N° d'ordre : 880

THÈSE

présentée devant

L'UNIVERSITE PAUL-SABATIER DE TOULOUSE
(SCIENCES)

pour obtenir

LE GRADE de DOCTEUR ES-SCIENCES MATHEMATIQUES
Mention : "Informatique Théorique"

par

François RODRIGUEZ

CONTRIBUTION A L'ETUDE DES OPERATIONS SUR LES LANGAGES ET DES OPERATEURS SUR LES FAMILLES DE LANGAGES

Soutenue le 20 juin 1979, devant la Commission d'examen

MM.	M.	NIVAT	Président
	L.	BOASSON	Examinateurs
	R.	BOOK	
	A.	BRUEL	
	R.	CORI	
	H.	GALLAIRE	
	J.	VIGNOLLE	

N° d'ordre : 880

THÈSE

présentée devant

L'UNIVERSITE PAUL-SABATIER DE TOULOUSE
(SCIENCES)

pour obtenir

LE GRADE de DOCTEUR ÈS-SCIENCES MATHEMATIQUES
Mention : "Informatique Théorique"

par

François RODRIGUEZ

CONTRIBUTION A L'ETUDE DES OPERATIONS SUR LES LANGAGES ET DES OPERATEURS SUR LES FAMILLES DE LANGAGES

Soutenue le 20 juin 1979, devant la Commission d'examen

MM.	M.	NIVAT	Président
	L.	BOASSON	Examinateurs
	R.	BOOK	
	A.	BRUEL	
	R.	CORI	
	H.	GALLAIRE	
	J.	VIGNOLLE	

UNIV.
LEIDEN
BIBL.

UNIVERSITE PAUL SABATIER

PRESIDENCE

M. MARTIN ..... Président
M. LARENG ..... 1er Vice-Président
Mlle GOUYON ..... 2ème Vice-Président

CORPS PROFESSORAL

ORDRE DES SCIENCES

HONORARIAT

M. BEDOS ..... Professeur honoraire
M. BLAIZOT ..... Doyen honoraire
M. CAPDECOMME .. Doyen honoraire, Recteur honoraire, Correspondant de l'Institut
M. COUCHET ..... Professeur honoraire
M. DUPOUY ..... Membre de l'Institut, Doyen honoraire, Directeur Général honoraire du C.N.R.S., Professeur honoraire
M. DURAND Emile .. Doyen honoraire, Professeur honoraire
M. GAUSSEN ..... Professeur honoraire, Correspondant de l'Institut
M. MARGULIS ..... Professeur honoraire
M. MASDUPUY ..... Professeur honoraire
M. MATHIS ..... Doyen honoraire
M. MIGNONAC ..... Professeur honoraire
M. MIQUEL ..... Professeur honoraire
M. MORQUER ..... Professeur honoraire, Correspondant de l'Institut
M. SECONDAT ..... Professeur honoraire
M. TEISSIE-SOLIER . Professeur honoraire
M. VANDEL ..... Professeur honoraire, Membre de l'Institut
M. DUPIN ..... Professeur honoraire
M. TRICHE ..... Professeur honoraire
M. PERRIER ..... Professeur honoraire

PROFESSEURS

M. CAPDECOMME .. Minéralogie, Correspondant de l'Institut
M. GALLAIS ..... Chimie, Membre de l'Institut
M. FERT ..... Physique
M. LESBRE ..... Chimie Organique
M. HURON ..... Mathématiques Appliquées
M. LEDOUX ..... Zoologie Appliquée
M. RIVALS ..... Agriculture
M. MATHIS ..... Chimie
M. ORLIAC ..... Géochimie et Minéralogie
M. LAFOURCADE .. Physique
M. ANGELIER ..... Zoologie
Mlle de FERRE ..... Botanique
M. FARRAN ..... Minéralogie et Géotechnique
M. LAUDET ..... Physique Théorique et Calcul Numérique
M. SERFATY ..... Physiologie et Ecophysiologie Animales
M. LAGASSE ..... Electrotechnique
M. BLANC ..... Physique Nucléaire
M. LEREDDE ..... Botanique
M. LELUBRE ..... Géologie

M. LALAGUE ..... Mathématiques Générales
M. BOUIGUE ..... Astronomie
M. ASSELINEAU .... Chimie Biologique
M. MAURET ..... Chimie Systématique
M. AGID ..... Physiologie
M. MONTANT ..... Cryptogamie
M. GAUTIER ..... Physique
M. CRUMEYROLLE . Mathématiques
M. GOURINARD .... Géologie
M. PULOU ..... Minéralogie
M. CAMBOU ..... Physique Spatiale
M. LACOSTE ..... Electrotechnique
M. THIBAULT ..... Mécanique Rationnelle et Appliquée
M. MASCART ..... Mathématiques
M. MEDIONI ..... Psychophysiologie
M. RAYNAUD P. .... Physiologie Animale
M. ZALTA ..... Chimie Biologique
M. SEVELY ..... Electrotechnique
M. POMMIEZ ..... Mathématiques
M. REY Paul ..... Biologie Végétale
M. COULOMB ..... Physique
M. TRINQUIER ..... Physique
M. MARONI ..... Chimie
M. BEETSCHEN .... Biologie Générale
M. DERACHE ..... Physiologie Animale
M. SATGE ..... Chimie Organique
M. LATTES ..... Chimie
M. VEDRENNE ..... Géophysique
M. DURAND-DELGA Géologie
M. CARRARA ..... Physique
M. MAHENC ..... Chimie
M. MIROUSE ..... Géologie
M. BITSCH ..... Zoologie
M. DEGEILH ..... Physique
M. MARTIN J.C. .... Génie Electrique
M. REY Gérard ..... Electronique, Electrotechnique, Automatique
M. SICARD ..... Biologie Génétique
M. SOUQUET ..... Géologie
M. TOUZE ..... Physiologie Végétale
M. FRASNAY ..... Mathématiques (Algèbre et Combinatoire)
M. CASSAGNAU .... Zoologie
M. CAUSSINUS ..... Mathématiques Appliquées (Statistiques Appliquées)
M. PESCIA ..... Physique
M. PICCA ..... Physique de l'Atmosphère
M. BAUDIERE ..... Botanique Fondamentale et Pyrénéenne
M. BARRANS ..... Chimie Physique Organique
M. POILBLANC ..... Chimie Minérale
M. PERENNOU ..... Informatique
M. ATTEIA ..... Mathématiques
M. CASTAN ..... Informatique
M. COLLETTE ..... Physique
M. REME ..... Mesures Physiques

PROFESSEURS SANS CHAIRE

M. MERIC ..... Mathématiques Appliquées
Mme LECAL ..... Zoologie
M. PILOD ..... Physique
M. LARROQUE ..... Physique
Mlle LAPEYRE ..... Mathématiques - Informatique
M. BERTRAND ..... Chimie
M. DESQ ..... Mathématiques
M. ROCARD ..... Electronique
M. GUERIN ..... Mathématiques
M. SCHNEIDER ..... Biologie Cellulaire
M. de LOTH ..... Chimie Physique
M. SAPORTE ..... Physique
M. THENOZ ..... Génie Civil
M. DURAND Ph. ..... Physique
M. FONTAN ..... Physique Nucléaire
M. BAUDRAS ..... Chimie Biologique
M. CALVET ..... Mécanique des Fluides
M. PAGANI ..... Physique
M. BEAUFILS ..... Informatique
M. BERTHELEMY ... Zoologie
M. TERJANIAN ..... Mathématiques
M. MORUCCI ..... Génie Biologique et Médical
M. BONEL ..... Chimie
M. SOTIROPOULOS . Chimie Organique
M. CUPPENS ..... Mathématiques
M. VERDIER ..... Physique
M. ETTINGER ..... Mathématiques
M. BONNET Louis ... Biologie
M. JOSSERAND .... Mesures Physiques
M. ROUTIE ..... Génie Chimique
M. COTTU ..... Génie Mécanique
M. HURAUX ..... Physique
Mme GERVAIS ..... Chimie Inorganique
M. BANCEL ..... Mathématiques
M. LOUARN ..... Génétique
M. BOUDET ..... Physiologie Végétale

PROFESSEUR ASSOCIE

M. GUMOWSKI ..... Mathématiques

MAITRES DE CONFERENCES

Mlle BARBANCE .... Mathématiques
M. GILLY ..... Génie Mécanique
M. MARAL ..... Physique
M. LEGRAND ..... Génie Civil
M. ABATUT ..... Electronique, Electrotechnique, Automatique
M. MAUSS ..... Mécanique
M. BETOURNE ..... Informatique
M. CAMPAN ..... Psychophysiologie
M. CLERC ..... Mécanique
M. GRIFONE ..... Mathématiques
M. LETAC ..... Mathématiques
M. COUOT ..... Mathématiques, Analyse Numérique
M. NGUYEN THANH VAN Mathématiques
M. TRAVERSE ..... Problèmes Chimiques de l'Energie
M. ALRAN ..... Génie Chimique
M. REY J. ..... Géologie Sédimentaire et Paléontologie
M. DARTIGUENAVE Chimie Minérale Moléculaire
M. HERAULT ..... Chimie
M. PRADINES ..... Mathématiques
M. GALINIER ..... Informatique
M. VIGNOLLE ..... Informatique
M. DEPARIS ..... Embryologie
M. CAVALIE ..... Physiologie Végétale
M. MASSOL ..... Chimie des Composés Organiques et Organominéraux d'intérêt biologique
M. HARTMANN .... Mécanique
M. ROUSSET ..... Chimie Appliquée (Matériaux)
M. TARDY ..... Sciences de la Terre et Aménagement

MAITRES DE CONFERENCES ASSOCIES

M. HURLEY ..... Physique Spatiale
M. SINGH ..... Automatique

ADMINISTRATION

M. PRINEAU ..... Secrétaire Général de l'Université

INSTITUT NATIONAL POLYTECHNIQUE DE TOULOUSE

PRESIDENCE

M. MONTEL ..... Président
M. MARTY ..... Vice-Président
M. ENJALBERT .... Vice-Président
M. ANDRE ..... Vice-Président
M. CONSTANT ..... Vice-Président

HONORARIAT

M. DIEHL ..... Professeur honoraire
M. ESCANDE ..... Professeur honoraire, Membre de l'Institut
M. BIREBENT ..... Professeur honoraire
M. CASTAGNETTO .. Maître de Conférences honoraire

PROFESSEURS

M. NOUGARO ..... Hydraulique Générale et Appliquée
M. GARDY ..... Génie Chimique
M. VOIGT ..... Chimie Minérale
M. MONTEL ..... Chimie Inorganique
M. THIRRIOT ..... Hydraulique
M. GRUAT ..... Hydraulique
M. BUGAREL ..... Génie Chimique
M. DAT ..... Hydraulique
M. MARTY ..... Electronique
M. ANGELINO ..... Génie Chimique

PROFESSEURS SANS CHAIRE

Mlle BERDUCOU .... Physiologie Végétale
M. HAMANT ..... Cythologie et Pathologie Végétales
M. FALLOT ..... Technologie Végétale
M. ENJALBERT .... Génie Chimique
M. HOFFMANN ..... Electronique
M. TRUCHASSON ... Hydraulique
M. DABOSI ..... Métallurgie et Réfractaires
M. LEFEUVRE ..... Electronique
M. CALMON ..... Chimie Agricole
M. GILOT ..... Génie Chimique
M. MATHIEU ..... Chimie Appliquée

MAITRES DE CONFERENCES

M. BAUDRAND ..... Electronique
M. TRANNOY ..... Electrotechnique
M. BOURGEAT ..... Pédologie
M. MATHEAU ..... Electronique
M. BAJON ..... Electronique
M. COUDERC ..... Génie Chimique
M. GOURDENNE ... Chimie
M. LENZI ..... Chimie Industrielle
M. MASBERNAT .... Hydraulique
M. TERRON ..... Zoologie
M. BUIS ..... Biologie Quantitative
M. CONSTANT ..... Chimie Minérale
M. COSTES ..... Electrotechnique
M. ECOCHARD ..... Agronomie
M. CANDAU ..... Zootechnie
M. LABAT ..... Ichtyologie
M. MORELIERE .... Electronique
M. GASET ..... Chimie Industrielle
M. BRUEL ..... Informatique
M. JAMET ..... Analyse Numérique

ADMINISTRATION

M. CRAMPES ..... Secrétaire Général

Liste établie au 1er Octobre 1977
pour l'année universitaire 1977-1978

à Nicole,

Je tiens à exprimer ma profonde reconnaissance à tous les membres du jury,

à Maurice NIVAT dont l'intérêt pour mes recherches et les conseils judicieux m'ont incité à poursuivre mes efforts. Je le remercie d'avoir bien voulu présider le jury de cette thèse,

à Luc BOASSON dont l'écoute attentive et exigeante m'a été des plus précieuse,

à R.V. BOOK qui a accepté de juger mon travail. Je suis particulièrement honoré de sa présence dans ce jury,

à André BRUEL, responsable de la Filière Informatique de l'E.N.S.E.E.I.H.T qui m'a toujours donné les plus grandes facilités pour mener à bien ce travail. Je tiens à le remercier une fois encore et lui dire combien j'y ai été sensible,

à Robert CORI qui a bien voulu s'intéresser à ce travail et faire partie de ce jury,

à Hervé GALLAIRE qui, en me proposant un sujet de seconde thèse, m'a fait découvrir un domaine de l'informatique particulièrement intéressant,

à Jean VIGNOLLE dont les encouragements incessants alliés à l'ambiance chaleureuse et dynamique qu'il a créé au sein de l'équipe Informatique Théorique de l'ERA 298 de Toulouse ont été déterminants dans la réalisation de ce travail.

Je tiens enfin à remercier tous mes amis de l'équipe d'Informatique Théorique de Toulouse et de l'Institut de Programmation de Paris qui n'ont cessé de m'encourager et me conseiller.

Je m'en voudrais d'oublier la gentillesse, la constante patience et l'excellence du travail dactylographique de Marie-Blandine VOLPATO ainsi que la qualité et la diligence apportées par Monsieur BERNARD à la réalisation typographique. Je les en remercie très vivement.

SOMMAIRE

INTRODUCTION	I à XII
CHAPITRE 1 : PRELIMINAIRES	1 à 16
A) Terminologie, notations	2 à 5
B) Rappels sur les transductions rationnelles, cônes rationnels et FAL (fidèles)	6 à 10
C) Rappels sur les opérations et opérateurs substitution, chevron et crochet	11 à 16
CHAPITRE 2 : OPERATIONS ET OPERATEURS	17 à 80
A) Introduction	18 à 20
B) Définitions préliminaires, propriétés élémentaires	21 à 27
C) Opérations régulières	28 à 34
D) Opérations indépendantes d'un opérateur	35 à 43
E) Opérations indépendantes de Rat	44 à 54
F) Opérations de degré $p \geq 0$	55 à 64
G) Opérations syntaxiques	65 à 70
H) FAL décomposables	71 à 80
CHAPITRE 3	81 à 126
A) L'opération crochet ouvert	82 à 99
B) Les opérations unaires crochet, chevron et crochet ouvert de degré i	100 à 119
C) Les opérations unaires crochet, chevron et crochet ouvert de degré infini	120 à 126

CHAPITRE 4 : RELATIONS RATIONNELLES N-AIRES 127 à 156

A) Introduction 128 à 129

B) Définitions et propriétés élémentaires 129 à 138

C) Lemme d'itération 139 à 140

D) Relations rationnelles particulières 140 à 150

E) Relations rationnelles et applications 151 à 156

CHAPITRE 5 : CHEVRON ET CROCHET OUVERT D'ORDRE n 157 à 170

A) Introduction 158 à 159

B) Propriétés générales 159 à 168

C) Applications 169 à 170

CHAPITRE 6 : INDEPENDANCE DES OPERATIONS 171 à 185

CHAPITRE 7 : LANGAGES VARIFIABLES ET OPERATIONS SYNTAXIQUES CLASSIQUES 187 à 204

A) Introduction 188 à 189

B) Langages vérifiables, propriétés élémentaires 189 à 192

C) Caractérisation et comparaison de diverses hiérarchies classiques 193 à 199

D) Sur diverses autres familles de langages 200 à 204

BIBLIOGRAPHIE 205 à 214

INTRODUCTION

La théorie des langages formels est née de la considération de problèmes linguistiques, essentiellement d'une tentative de caractérisation mathématique des langues naturelles par leur processus de génération, appelé grammaire. Dès 1956, N. Chomsky définissait quatre types de grammaires [Chom1,2] dont le premier est celui des grammaires à structure de phrase qui engendrent la classe des ensembles récursivement énumérables [Rog]. Les trois types suivants en sont des cas particuliers : grammaires à contexte lié (ou "context sensitive") qui engendrent les ensembles récursifs, grammaires à contexte libre qui engendrent les langages que nous appellerons algébriques (ou "context free") et enfin les grammaires régulières associées aux langages rationnels (ou "réguliers").

Par ailleurs, les circuits séquentiels constituaient un élément de base des ordinateurs. Le lien entre les circuits séquentiels et les langages réguliers fut rapidement mis à jour et exploité de façon intensive : la théorie des automates fut d'abord celle des automates finis et la "switching theory". D'autre part les informaticiens cherchaient à élaborer des langages de programmation pratiques et efficients. Les langages rationnels étant insuffisants pour modéliser de tels langages, on eut recours aux langages algébriques et aux automates à pile qui les reconnaissent. Ces modèles ont par exemple présidé à la conception du langage ALGOL 60, ce qui explique la dénomination "Algol-like languages" utilisée pour désigner les langages algébriques dans la littérature anglo-saxonne.

L'étude systématique des langages algébriques et des automates à pile et notamment la considération de sous-classes particulièrement intéressantes (langages déterministes [GG4], grammaires LR(k) [Knu],...) permit ainsi la simplification des processus d'analyse syntaxique et, de façon générale, d'aborder la compilation des langages évolués avec beaucoup plus de rigueur.

Ces faits, seuls, suffiraient à justifier le développement de cette branche de la théorie. Cependant l'aspect prépondérant des langages

algébriques fut souligné de façon nettement plus remarquable par les travaux de Schützenberger [Shutz1,2] qui inscrivait la formalisation de cette théorie dans un domaine classique des mathématiques : les séries formelles en variables non commutatives (cf [Flie]). La rencontre du linguiste Chomsky et du mathématicien Schützenberger devait donc définitivement consacrer la théorie des langages algébriques [CS].

Il est important de souligner, qu'outre les progrès qui en résultèrent en théorie des langages de programmation ou en théorie des séries formelles [Jac1,2], les langages algébriques ont permis d'obtenir des résultats importants en sémantique (théorie des schémas de Programme [Cour,CR]) et même en théorie de la complexité [BNP,BCON,Gr9]. De plus ces divers aspects se complètent souvent réciproquement. Ainsi les langages d'interprétation de certains schémas de programme induisent des propriétés au niveau des schémas et, inversement, l'importance de certains types de schémas a déterminé l'étude de sous-familles de langages algébriques et fait apparaître des phénomènes ignorés jusque là [Aut2].

L'importance et la diversité des résultats acquis en théorie des langages algébriques et leurs multiples conséquences dans des domaines apparemment disjoints, ne résolvent pas, cependant, un certain nombre de problèmes inhérents à la compilation des langages évolués. Les exemples les plus simples sont ceux des déclarations (de type) et des étiquettes. La syntaxe d'un langage tel qu'ALGOL est algébrique "presque partout" : la gestion des déclarations et des étiquettes ne peut pas être réalisée par un automate à pile. D'autre part, les familles de langages très générales définies initialement par Chomsky, et qui contiennent les langages algébriques, sont associées à des classes d'automates inutilisables, raisonnablement, en compilation.

Un modèle complet et performant permettant de résoudre ces problèmes a été alors introduit en 1967 par Ginsburg, Greibach et Harrison sous le nom de "one-way stack automata" [GGH1,2]. Ces automates unilatères sont munis d'un ruban mémoire sur lequel l'écriture et l'effacement se font uniquement en "sommet" de mémoire (comme pour les automates à pile), mais dont on peut également consulter le contenu. Ces automates permettent une simulation complète des processus de compilation. En fait, dans ce cas, les automates ont

une gestion mémoire très particulière : le ruban mémoire comporte une partie fixe où sont stockées les déclarations que l'on viendra vérifier au cours de la compilation, l'autre partie du ruban fonctionnant comme une pile pour l'examen des sous-ensembles du programme qui sont algébriques.

Une sous-classe des "one-way stack automata" s'est révélée particulièrement importante : les accepteurs vérificateurs ("checking automata") qui écrivent un mot en mémoire puis en consultent le contenu sans plus pouvoir le modifier. La famille des langages acceptés par ces automates, appelés langages vérifiables [Gr5,Rod6], présente en effet de nombreuses analogies structurelles avec la famille des langages algébriques. De plus il s'agit d'une famille de taille "raisonnable" qui contient de nombreux langages algébriques (linéaires, 1-compteur) ainsi que des langages tels que $L_1 = \{wcw | w \in \{a,b\}^*\}$ et $L = \{a^n b^n c^n | n \in \mathbb{N}\}$. Nous verrons que le langage L_1, en particulier, se révèle avoir d'intéressantes propriétés, analogues aux langages linéaires. Signalons en outre que, récemment, un calcul automatique de la définition formelle des fonctions calculées par certains programmes récursifs conduit naturellement à des structures voisines de celles du langage L_1 [Greu].

L'introduction de ces nouveaux automates et des familles de langages associées devait influencer la théorie de façon déterminante. De même que la famille des langages algébriques, celle des langages vérifiables et celle des langages reconnus par les "one-way stack automata" sont fermées par homomorphisme, homomorphisme inverse, intersection avec les langages rationnels et étoile (monoïde engendré). De plus chaque langage de ces familles s'exprime à partir d'un seul par application d'un nombre fini de fois les opérations précédentes. Ces similitudes de propriétés devaient donner naissance à la théorie des "Abstract families of languages" (AFL en abrégé, famille agréable de langage ou FAL dans la littérature française) et à son équivalent en termes de machines les "abstract families of acceptors" [GG1,2,...]. Ces deux théories permettent d'unifier l'ensemble des études faites jusque là, d'établir simplement des résultats connus et d'éclairer de façon nouvelle et particulièrement fructueuse les rapports entre les divers types d'automates et les classes de langages qu'ils acceptent. Un grand nombre de problèmes nouveaux furent ainsi mis à jour et assez souvent résolus.

De façon totalement indépendante à partir de la notion de transduction introduite par Elgot et Mezei [EM], M. Nivat développait une théorie purement algébrique des transductions rationnelles dont il donnait à la fois les propriétés et une caractérisation particulièrement manipulable en termes de morphismes alphabétiques [Ni1]. Cette théorie est un outil privilégié pour étudier les propriétés de fermeture par homomorphisme, homomorphisme inverse et intersection avec les langages réguliers.

De plus les transductions rationnelles permettent de réécrire la théorie des FAL d'une façon beaucoup plus claire et agréable en utilisant la notion de cône rationnel [BO2] et donc de retrouver simplement les résultats de [GG1,2,...]. Notons enfin que l'introduction naturelle de la relation "domine par transduction rationnelle" permet de rendre compte précisément des classifications classiques et notamment dans la famille des langages algébriques. Elle peut donc être utilisée fructueusement comme indicateur de la complexité de ces langages : les travaux de Boasson sur les paires itérantes en sont un des meilleurs exemples [BONi1,BO2].

Cependant la notion de cône rationnel comme celle de FAL ne suffit pas pour rendre compte de la structure algébrique précise de certaines classes de langages et l'introduction d'opérations plus spécifiques sur les langages s'est avérée nécessaire. Ainsi dans [Ni2], Nivat utilise la substitution pour caractériser les langages quasirationnels. De même les langages récursivement énumérables ont été démontrés être l'image par homomorphisme de l'intersection de deux langages à 2-compteurs [Mi] puis de l'intersection de deux langages appartenant à la FAL engendrée par $\{a^n b^n \mid n \in \mathbb{N}\}$ [HH]. Inversement Greibach montrait dans [Gr2] que si un cône rationnel n'est pas fermé par substitution, alors l'opérateur substitution permet la construction d'une hiérarchie infinie de cônes rationnels. L'aspect général de cette propriété, c'est-à-dire son indépendance vis-à-vis de la famille de langages considérée a conduit Greibach à développer une étude systématique des opérations binaires ayant de telles propriétés [Gr3]. La théorie des opérations binaires syntaxiques (ou syntactiques) ainsi construite rendait alors parfaitement compte des phénomènes qui déterminent la non principalité de nombreuses familles de langages classiques. En outre elle permettait d'établir simplement de nouvelles caractérisations purement algébriques de ces familles.

La théorie des opérations binaires syntaxiques se révélait pourtant vite insuffisante, tandis que se confirmait l'intérêt d'utiliser de nouvelles opérations pour parfaire l'étude de nombreuses familles de langages. Ainsi dans [BCN] les auteurs caractérisent les langages ultralinéaires, quasirationnels et fortement quasirationnels à l'aide de deux opérations unaires appelées chevron et crochet qui leur permettent d'établir la non principalité de ces familles. Les résultats obtenus dans [BCN] sont similaires à ceux de [Gr3] pour les opérations binaires syntaxiques mais, comme le souligent Boasson, Crestin et Nivat ces résultats "ne peuvent être simplement déduits des résultats très généraux" de [Gr3]. En effet d'une part le chevron et le crochet sont des opérations unaires, d'autre part l'opérateur crochet est idempotent, i.e. il suffit de l'appliquer une fois sur une famille de langages pour obtenir sa fermeture par crochet, contrairement à ce qui se passe pour les opérations binaires syntaxiques de [Gr3]. Lorsque nous avons introduit dans [Rod4] une nouvelle opération unaire, appelée crochet ouvert, pour étudier la famille deslangages vérifiables et notamment le cône rationnel engendré par $L_1 = \{wcw|w \in \{a,b\}^*\}$, nous avons été conduit aux mêmes conclusions. Cependant, cette opération a de surprenantes analogies tant avec le chevron et le crochet qu'avec les opérations binaires syntaxiques de [Gr3]. Il semblait donc que l'arité des opérations soit indépendante de ces phénomènes de même que l'idempotence éventuelle des opérateurs. Ces constatations -et contradictions apparentes- nous ont alors conduit à rechercher les propriétés communes à toutes ces opérations et notamment celles qui déterminent cette similitude de résultats, au demeurant établis par des procédés différents.

Notre travail se décompose en fait en deux parties : d'une part la construction d'une théorie des opérations et des opérateurs, d'autre part son utilisation systématique à l'étude d'opérations classiques et surtout celle du crochet ouvert et de diverses généralisations que nous utilisons pour étudier les langages vérifiables.

L'originalité de la première partie par rapport à la théorie de Greibach dans [Gr3] réside en deux points essentiels. Tout d'abord les opérations considérées sont d'une arité quelconque, $n \geq 1$. S'il était clair que de nombreuses propriétés des opérations binaires syntaxiques de [Gr3] se généralisaient simplement à des opérations d'arité quelconque, cela n'était

absolument pas évident pour d'autres : citons par exemple l'existence d'une FAL maximale dans les FAL principales fermées par des opérations unaires. En second lieu les classes d'opérations que nous étudions sont beaucoup plus générales que celles de [Gr3] puisqu'elles englobent aussi bien l'opération étoile (monoïde engendré) que le crochet ou les opérations de [Gr3]. Nous introduisons, à ce niveau, deux concepts nouveaux. D'une part les opérations régulières qui préservent la relation "domine par transduction rationnelle", d'autre part les opérations indépendantes d'un opérateur. Cette dernière notion est plus forte que l'indépendance des opérations étudiée dans [GH,BG] et deux cas particuliers sont fondamentaux : les opérations indépendantes de Rat, l'opérateur de fermeture rationnelle, et les opérations indépendantes de l'opérateur auquel elles sont associées.

La théorie ainsi développée nous permet à la fois de présenter de façon unifiée les propriétés remarquables de nombreuses opérations considérées en théorie des langages formels, de simplifier plusieurs preuves de résultats connus, d'étendre un certain nombre de résultats classiques (langages linéaires, langages bornés,...) et d'établir enfin de nouveaux théorèmes sur des familles de langages comme $\mathscr{L}_{BNP}$ et $\mathscr{L}_{DUP}$ [BNP,BG].

Dans une deuxième partie, nous utilisons les résultats précédents pour étudier l'opération crochet ouvert et des généralisations des opérations chevron, crochet et crochet ouvert. La régularité de ces opérations nous permet d'établir en particulier qu'elles préservent une propriété, fondamentale du point de vue de la complexité, appelée "effacement" [Lat2]. Ceci nous permet d'exhiber de nombreux langages, algébriques ou non, qui sont "effaçables". Les opérations étudiées nous permettent enfin de donner des caractérisations algébriques simples de plusieurs familles de langages définies soit par des automates, soit par des grammaires généralement complexes. Ces caractérisations nous permettent alors de comparer sans peine diverses hiérarchies de langages vérifiables.

Indiquons enfin que les méthodes que nous avons présentées se révèlent, à ce jour, impuissantes pour résoudre de nombreux problèmes ouverts de la théorie. Citons comme premier exemple la célèbre, et déjà ancienne, question de savoir si la FAL des langages algébriques non générateurs est

principale ou non. Le second problème, plus récent, que nous abordons d'ailleurs dans ce travail dans des cas particuliers, est celui des propriétés des limites de suites d'opérations sur les langages. Ce problème est équivalent à celui toujours non résolu de la caractérisation de l'intersection de deux familles de langages. Toutefois nos méthodes permettent à la fois de préciser ces conjectures et d'établir, dans le deuxième exemple, que certaines propriétés sont préservées par le passage à la limite.

De façon plus précise, notre travail a été divisé en sept chapitres :

- Le chapitre 1 est un chapitre préliminaire où nous rappelons la terminologie et les propriétés classiques de la théorie des langages que nous utilisons dans la suite : transductions rationnelles, cônes rationnels (fidèles) et FAL (fidèles). Les rapports entre ces notions sont rappelés de façon précise. Ce chapitre se termine par un exposé des propriétés caractéristiques des opérations substitution, chevron et crochet établies notamment dans [Gr2] et [BCN]. Ces exemples seront utilisés systématiquement dans les chapitres suivants.

- Le deuxième chapitre constitue l'étude générale des opérations sur les langages et des opérateurs sur les familles de langages. Il se décompose en huit parties :

A) Une introduction où nous développons les arguments qui nous ont conduit à cette étude.

B) Un exposé des définitions générales (opérations régulières, opérateurs simples, coniques, associés à une opération) et l'énoncé de deux lemmes généraux utilisés par la suite.

C) Cette partie est consacrée à l'étude des opérations régulières, i.e. qui préservent la relation "domine par transduction rationnelle". Nous montrons que si T est un opérateur régulier, c'est-à-dire associé à une opération régulière alors T préserve la principalité des cônes rationnels (Théorème 2.1). Il en résulte plusieurs corollaires qui nous permettent de

retrouver simplement des résultats de [GG2] sur la substitution et de généraliser le lemme syntaxique de [Beau1]. Le théorème 2.1 bis donne une condition nécessaire et suffisante pour qu'un opérateur soit régulier.

D) La notion d'opération indépendante d'un opérateur est introduite dans cette partie. Elle nous permet d'établir un théorème d'où découle la non principalité de certains cônes rationnels (théorème 2.2). Ce résultat englobe à la fois ceux de Greibach dans [Gr3] et ceux de [BCN] pour le chevron et le crochet. Le théorème 2.3 donne une condition nécessaire et suffisante pour que les transformés par un opérateur T_1 soient fermés par un opérateur T_2, si T_1 est indépendant de T_2. Cette partie s'achève par une condition nécessaire (et suffisante) pour qu'une opération (régulière) soit indépendante d'un opérateur (lemme 2.7). Nous retrouvons ainsi un résultat de [Gr4] sur les linéaires.

E) Les opérations indépendantes de Rat, l'opérateur de fermeture rationnelle, ont une importance primordiale. De la partie D nous déduisons en effet que si $\mathscr{L}$ est un cône rationnel tel que la FAL $\mathscr{F}(\mathscr{L})$ engendrée par $\mathscr{L}$ est fermée pour une telle opération alors $\mathscr{L}$ est lui même une FAL (corollaire 2.12). Ce résultat généralise donc à une vaste classe d'opérations le théorème 2.1 de [Gr3] sur les cônes rationnels "uniformly star closed".

Le théorème 2.4 donne une condition suffisante pour que la propriété suivante soit vérifiée par un cône rationnel $\mathscr{L}$: $\mathscr{L}$ est principal si et seulement si $\mathscr{F}(\mathscr{L})$ est principal. Ce théorème nous permet de simplifier plusieurs preuves et de préciser les relations entre les langages ultralinéaires et fortement quasi rationnels et les FAL engendrées correspondantes (propositions 2.1 et 2.2). Nous concluons alors que la FAL engendrée par le plus petit cône rationnel fermé par chevron n'est pas la FAL engendrée par les langages algébriques bornés [BCN].

Nous étudions alors plus particulièrement deux applications des résultats précédents à des familles classiques. On montre que Rat, la famille des langages rationnels est la seule FAL fermée par une opération indépendante de Rat contenue dans la FAL engendrée par les langages bornés (proposition 2.3). De même Rat est le seul cône rationnel fermé par produit et par une opération

indépendante de Rat contenu dans la FAL des langages linéaires (proposition 2.4). On obtiendra d'ailleurs un résultat analogue pour la FAL engendrée par $L_1 = \{wcw | w \in \{a,b\}^*\}$ au chapitre 3 (corollaire 3.15).

F) Une opération de degré 0 est indépendante de l'opérateur qui lui est associé. La définition que nous utilisons est en fait plus restrictive, mais l'ensemble des résultats de cette partie reste vraie pour les opérations indépendantes de leur opérateur. Une opération est de degré $p > 0$ si elle vérifie une propriété un peu plus complexe, nécessaire pour traiter le cas d'opérations comme l'insertion [Gr3] qui n'est pas de degré 0. Nous montrons alors que si $\mathcal{L}$ est un cône rationnel non fermé par une opération de degré $p \geq 0$, alors la fermeture de $\mathcal{L}$ pour cette opération est non principale (théorème 2.8). Ceci généralise le théorème 2.2 de [Gr3] en englobant notamment le cas de l'opération chevron. Introduisant alors la notion d'opération strictement de degré p, nous donnons un résultat qui répond de façon générale à un problème ouvert de Ginsburg [Gi2,p 242] (théorème 2.9).

G) Reprenant la terminologie de [Gr3] nous appelons opération syntaxique (ou syntactique) une opération de degré $p \geq 0$ indépendante de Rat. Des parties E et F on déduit alors simplement des propriétés analogues à celles de [Gr3] (théorème 2.10, corollaire 2.21). Cette partie se termine par l'étude d'une opération équivalente à l'insertion, qui nous permet de répondre à un problème ouvert de [Gr3,p 38] et nous conduit à poser une nouvelle conjecture.

H) Une FAL $\mathcal{L}$ est décomposable s'il existe deux FAL incomparables $\mathcal{L}_1$ et $\mathcal{L}_2$ telles que $\mathcal{L} = \mathcal{F}(\mathcal{L}_1 \cup \mathcal{L}_2)$. Nous montrons que toute FAL $\mathcal{L}$ principale fermée par une opération syntaxique n-aire, $n \geq 1$, est indécomposable et recouvre donc une FAL maximale qui est la famille des non générateurs de $\mathcal{L}$ (théorème 2.13.). De plus ce résultat reste vrai dans le cas des FAL fidèles fermées par homomorphisme à effacement linéaire ce qui répond à une question de [Gr3,p 46]. Ceci nous permet de conclure que la FAL fidèle $\mathcal{L}_{DUP}$ [BG] recouvre une FAL maximale. Les cônes rationnels fidèles fermés par produit et homomorphisme à effacement linéaire vérifient également des propriétés analogues au théorème 2.13. Nous en déduisons que $\mathcal{L}_{BNP}$ [BNP] recouvre un cône rationnel fidèle maximal.

Nous exhibons enfin un nouveau langage $L_1 = \{wcw | w \in \{a,b\}^*\}$, qui est tel que la FAL engendrée est indécomposable mais non fermée par une opération syntaxique.

- Le chapitre 3 est consacré à l'étude de plusieurs opérations unaires. Il se décompose en trois parties :

A) La première est consacrée à l'étude de l'opération crochet ouvert définie dans [Rod4]. Cette opération associe à tout langage L le langage $CO(L) = \{w\ell w | w \in \{a,b\}^*, \ell \in L\}$. L'utilisation systématique des résultats du chapitre 2 nous permet de démontrer, comme le crochet ouvert est régulier, que cette opération préserve l'effacement. Autrement dit si L est un langage tel que $\mathcal{E}^f(L) = \mathcal{E}(L)$ alors $\mathcal{E}^f(CO(L)) = \mathcal{E}(CO(L))$ (théorème 3.2). Nous remarquons qu'il en est de même pour l'opération crochet de [BCN] ce qui nous permet d'exhiber une infinité de langages effaçables nouveaux. Utilisant une caractérisation analogue à celle de [BCN] pour le crochet (théorème 3.1) et un résultat de [Lat4] nous montrons alors successivement que CO est une opération syntaxique de degré 0 (théorème 3.6) et que le produit de deux langages est indépendant de l'opérateur crochet ouvert (théorème 3.7). Nous déduisons également de cette étude l'incomparabilité des cônes rationnels engendrés respectivement par $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et $L_1 = \{wcw | w \in \{a,b\}^*\}$ (corollaire 3.12). Enfin après avoir caractérisé $\mathcal{E}^U_{CO}(Rat)$ le plus petit cône rationnel fermé par crochet ouvert nous montrons que Rat est le seul cône rationnel fermé par produit et par une opération indépendante de Rat contenu dans la FAL engendrée par $\mathcal{E}^U_{CO}(Rat)$.

B) Nous étudions ici des généralisations unaires du chevron, crochet et crochet ouvert que nous appelons chevron, crochet et crochet ouvert de degré i. De façon analogue à la partie A nous montrons que ces opérations sont régulières et que le crochet et le crochet ouvert de degré i préservent l'effacement (corollaire 3.17). Utilisant alors le fait que le produit de deux langages est indépendant des opérateurs Ch_i, C_i et CO_i (théorème 3.10) nous montrons que les trois opérations sont syntaxiques de degré 0 si $i \geq 2$ (théorème 3.11).

C) Nous associons aux suites d'opérations étudiées en B une opération limite. Nous définissons ainsi le crochet de degré infini par $C_\infty(L) = \bigcup_{i \geq 1} C_i(L)$ et le crochet ouvert de degré infini $CO_\infty(L) = \bigcup_{i \geq 1} CO_i(L)$. Nous remarquons que l'indépendance vis-à-vis de Rat est préservée alors que les autres propriétés du crochet (resp. crochet ouvert) de degré i ne le sont pas. En effet l'application répétée trois fois de ces opérateurs conduit à la fermeture. Nous identifions pour cela les plus petits cônes rationnels fidèles fermés par C_∞ et CO_∞ respectivement à $\mathcal{L}_{BNP}$ le plus petit cône rationnel fidèle fermé par intersection contenant les linéaires [BNP] et à $\mathcal{L}_{DUP}$ le plus petit cône rationnel fidèle fermé par intersection contenant $L_1 = \{wcw | w \in \{a,b\}^*\}$ [BG].

- Le chapitre 4 est consacré à l'étude des relations rationnelles n-aires dont nous nous servons tout au long du chapitre 5. Les résultats énoncés dans les parties B et C sont souvent connus [Bers,Ni1,Ei] où sont obtenus par des généralisations simples du cas n = 2. Nous ne reprenons donc pas la plupart des preuves. Dans la partie D nous étudions plus particulièrement certaines relations spécifiques en vue du chapitre 5. Dans la partie E nous considérons les applications rationnaire (n-1)-aires associées et constatons qu'en général la fermeture par transduction rationnelle n'implique pas la fermeture pour ces applications.

- Dans le chapitre 5 nous étudions deux familles d'opérations n-aires, $n \geq 1$, appelées chevron et crochet ouvert d'ordre n. Grâce à un théorème de caractérisation analogue au théorème 3.1 qui utilise des relations rationnelles n-aires (théorème 5.1) nous établissons le lien entre l'opération produit des langages et les opérateurs crochet ouvert d'ordre n (proposition 5.2). Ceci nous permet de conclure que le chevron et le crochet ouvert d'ordre n sont des opérations syntaxiques de degré 0.

Nous utilisons alors ces opérations pour donner une nouvelle caractérisation des "equal matrix languages" d'ordre n définis dans [Si1]. Nous établissons ainsi que ces familles de langages sont linéairement effaçables.

- Dans le chapitre 6 nous étudions l'indépendance des opérations étudiées précédemment vis-à-vis des opérations étoile, substitution et "homomorphic replication" [Gr3]. La définition de l'indépendance que nous utilisons est la suivante : t est indépendante de t' si pour tout cône rationnel $\mathcal{L}$ la condition $\mathcal{E}^U_{t'}(\mathcal{L}) \subseteq \mathcal{E}^U_t(\mathcal{L})$ implique $\mathcal{E}^U_{t'}(\mathcal{L}) \subseteq \mathcal{L}$. Nous montrons que toute opération indépendante de Rat est indépendante de l'étoile propre et une réciproque dans un cas particulier (corollaires 6.1 et 6.2). La proposition 6.1 sur le chevron, crochet et crochet ouvert résume en fait des résultats établis dans les chapitres précédents. De manière analogue à [Gr6] nous montrons alors que "l'homomorphic replication" est indépendante des opérations chevron, crochet et crochet ouvert (théorème 6.1) tandis que la réciproque est fausse. De même (théorème 6.2) la substitution est indépendante de ces trois opérations. Afin d'étudier la réciproque nous rappelons la définition des langages fortement k-itératifs [Gr6] et montrons -grâce au lemme d'itération sur les relations rationnelles du chapitre 4- que le crochet, le chevron et le crochet ouvert préservent cette propriété. Le théorème 6.3 établit alors que ces trois opérations sont indépendantes de la substitution dans un cône rationnel fortement 1-itératif. Dans le cas général cette réciproque est fausse.

- Le chapitre 7 utilise l'ensemble des résultats précédents pour étudier diverses familles de langages vérifiables. Nous commençons par rappeler la définition des accepteurs vérificateurs et celle des familles Cons_k et Mouv_k des langages reconnus par des accepteurs vérificateurs à consultation bornée ("finite visit") et à mouvements bornés ("finite reversal") [Rod1,6,Gr8]. Nous comparons alors ces familles entre elles et aux familles $\mathcal{D}_n$ des "equal matrix languages" d'ordre n [Si1]. Ce chapitre s'achève par la comparaison de ces familles aux cônes rationnels (fidèles) engendrés par des langages définis dans [BGW].

CHAPITRE 1

PRÉLIMINAIRES

A) TERMINOLOGIE, NOTATIONS

B) RAPPELS SUR LES TRANSDUCTIONS RATIONNELLES, CONES RATIONNELS ET FAL (FIDELES)

C) RAPPELS SUR LES OPERATIONS ET OPERATEURS SUBSTITUTION, CHEVRON ET CROCHET

A) TERMINOLOGIE, NOTATIONS

Etant donné un ensemble infini dénombrable $\sum$, appelé alphabet universel, si X est un sous-ensemble fini de $\sum$ nous désignerons par X^* le monoïde libre engendré par X pour l'opération de concaténation. Nous noterons ε_X (ou ε s'il ne peut y avoir de confusion) l'élément neutre de X^* ("mot vide"). Les éléments de X sont appelés des lettres, ceux de X^* des mots. Toute partie du monoïde libre X^* est un langage sur X. Pour tout langage L, X_L désignera la plus petite partie de $\sum$ telle que L est inclus dans X_L^*.

Dans l'ensemble des parties de X^*, nous considèrerons les opérations suivantes :

- L'union, notée $\cup$, qui est l'union ensembliste.

- Le produit, noté $.$, qui est induit par la concaténation dans X^*, défini par :

$$A.B = \{xy \mid x \in A,\ y \in B\}.$$

- L'étoile, notée $*$, qui est le passage au sous-monoïde engendré, définie par :

$$A^* = \bigcup_{n \geq 0} A^n \text{ avec } A^0 = \{\varepsilon\} \text{ et } A^{n+1} = A^n.A = A.A^n,\ \forall n \geq 0.$$

- L'étoile propre, notée $+$, définie par :

$$A^+ = \bigcup_{n \geq 1} A^n$$

- L'intersection, notée $\cap$, qui est l'intersection ensembliste.

Une famille de langages $\mathcal{L}$ est un ensemble de parties de $\sum^*$

tel que :

1) Pour tout $L \in \mathcal{L}$, il existe une partie finie X de $\sum$ telle que $L \subseteq X^*$.

2) $\mathcal{L}$ contient au moins un langage non vide.

Un homomorphisme φ de X^* dans Y^* est alphabétique (resp. strictement alphabétique, non effaçant) si $\varphi(X)$ est inclus dans $Y \cup \{\varepsilon\}$ (resp. Y, Y^+). φ est à effacement linéaire s'il existe $k \in \mathbb{N}$ tel que pour tout w, $|w| \geq k$ on a $|w| \leq k|\varphi(w)|$.

Une famille de langages $\mathcal{L}$ est recopiable si pour tout isomorphisme alphabétique φ et tout langage $L \in \mathcal{L}$, $\varphi(L)$ appartient à $\mathcal{L}$. Le langage $\varphi(L)$ sera appelé une recopie de L. Les familles de langages que nous considèrerons par la suite seront le plus souvent recopiables.

Définissons maintenant deux familles fondamentales :

Définition 1.1

Si X est un alphabet fini, Fin(X) désignera la famille des parties finies de X^*.

Rat(X), la famille des langages rationnels sur X^*, est la plus petite famille de langages de X^* qui contient Fin(X) et est fermée par union, produit et étoile.

Nous nous intéresserons souvent, par la suite, à des langages L propres, i.e. tels que $\varepsilon \notin L$. Nous définirons ainsi :

Définition 1.2

$$Fin_0(X) = \{L \in Fin(X) | \varepsilon \notin L\}$$

$$Rat_0(X) = \{L \in Rat(X) | \varepsilon \notin L\}$$

On démontre sans difficulté que $Rat_0(X)$ est la plus petite famille de langages qui contient $Fin_0(X)$ et qui est fermée par union, produit et étoile propre.

La famille Rat(X) peut être également caractérisée par les automates finis :

Définition 1.3

Un automate fini est un 5-uple $A = (Q,X,\delta,q_0,F)$, où :

- Q est un ensemble fini non vide d'états,
- $q_0 \in Q$ est l'état initial,
- $F \subseteq Q$ est l'ensemble des états finals,
- δ est une application de $Q \times X$ dans Q.

On étend, de façon classique, δ en une application de $Q \times X^*$ dans Q par :

- $\forall q \in Q,\ \delta(q,\varepsilon) = q$,
- $\forall a \in X,\ w \in X^*,\ q \in Q,\ \delta(q,wa) = \delta(\delta(q,w)a)$.

La langage reconnu par l'automate fini A est l'ensemble

$$L(A) = \{w \in X^* \mid \delta(q_0,w) \in F\}.$$

On a alors le théorème :

Théorème 1.1 de Kleene

Un langage L est rationnel si et seulement si il est reconnu par un automate fini.

Définition 1.4

Soient $X_1,\ldots,X_n$ des alphabets finis. Une partie A de $X_1^* \times X_2^* \times \ldots \times X_n^*$ sera dite rationnelle si A appartient à la fermeture de Kleene $(\cup,.,*)$ des parties finies de $X_1^* \times \ldots \times X_n^*$. Nous noterons $Rat_n(X_1,\ldots,X_n)$ l'ensemble des parties rationnelles de $X_1^* \times \ldots \times X_n^*$.

Nous écrirons Rat(X) pour $Rat_1(X)$ et $Rat_n(X)$ pour $Rat_n(X_1,\ldots,X_n)$ où $X_1 = X_2 = \ldots = X_n = X$.

Définition 1.5

Tout langage L sur X sera dit algébrique si c'est un langage "context-free" au sens de Ginsburg [Gi1] ou Eilenberg [Ei1].

B) RAPPELS SUR LES TRANSDUCTIONS RATIONNELLES, CONES RATIONNELS ET FAL (FIDELES)

Etant donné deux alphabets finis X et Y, on appelle transduction [EM] toute partie $\hat{\tau}$ de $X^* \times Y^*$. A une telle partie on associe deux applications -appelées encore transduction et transduction inverse- notée par τ et τ^{-1}, respectivement de X^* dans 2^{Y^*} et de Y^* dans 2^{X^*}, définies par :

$$\tau(f) = \{g \in Y^* | (f,g) \in \hat{\tau}\}, \forall f \in X^*$$

$$\tau^{-1}(g) = \{f \in X^* | (f,g) \in \hat{\tau}\}, \forall g \in Y^*$$

On étend alors de façon canonique τ (resp. τ^{-1}) en une application de 2^{X^*} dans 2^{Y^*} (resp. de 2^{Y^*} dans 2^{X^*}).

Une transduction τ de X^* dans Y^* est dite rationnelle si et seulement si $\hat{\tau} \in Rat_2(X,Y)$ [Ni1].

Une transduction est fidèle si et seulement si pour tout mot $g \in Y^*$, l'ensemble $\{x|(x,y) \in \hat{\tau}\}$ est fini.

Une transduction est continue si et seulement si $(z,\varepsilon) \in \hat{\tau}$ implique $z = \varepsilon$.

Les transductions rationnelles seront les seules utilisées dans la suite de ce travail. Elles sont rendues manipulables par les quatre théorèmes suivants :

Théorème 1.2 de Nivat [Ni1]

Une transduction τ de X^* dans Y^* est rationnelle si et seulement si il existe un alphabet Z, un langage rationnel $R \subseteq Z^*$ et deux homomorphismes alphabétiques φ et ψ de Z^* dans X^* et Y^* respectivement tels que

$$\hat{\tau} = \{(\varphi w, \psi w) | w \in R\}.$$

Théorème 1.3 [BoNi1]

Une transduction τ de X^* dans Y^* est rationnelle fidèle si et seulement si il existe un alphabet Z, un langage rationnel $R \subseteq Z^*$ et deux homomorphismes alphabétiques φ et ψ de Z^* dans X^* et Y^* respectivement, tels que :

1) $\hat{\tau} = \{(\varphi(w), \psi(w)) | w \in R\}$.

2) Il existe un entier k tel que $R \cap Z^* Z_0^k Z^* = \emptyset$ pour $Z_0 = \{z \in Z | \psi(z) = \varepsilon\}$.

Théorème 1.4 [BoNi1]

Une transduction τ de X^* dans Y^* est rationnelle continue si et seulement si il existe un alphabet Z, un langage rationnel $R \subseteq Z^*$ et deux homomorphismes φ et ψ de Z^* dans X^* et Y^* respectivement avec ψ strictement alphabétique, tels que :

$$\hat{\tau} = \{(\varphi(w), \psi(w)) | w \in R\}.$$

Théorème 1.5 [Bo1]

Si τ est une transduction rationnelle fidèle de X^* dans Y^* et $L \subseteq X^*$ un langage contenant le mot vide, il existe une transduction rationnelle fidèle et continue τ' telle que $\tau(L) = \tau'(L)$.

Rappelons enfin la propriété fondamentale de composition :

Théorème 1.6 [EM]

Soient τ_1 et τ_2 deux transductions rationnelles (resp. rationnelles fidèles, rationnelles fidèles et continues), la première de X^* dans Y^* et la seconde de Y^* dans T^*. Alors la transduction $\tau = \tau_2 \circ \tau_1$ de X^* dans T^* est rationnelle (resp. rationnelle fidèle, rationnelle fidèle et continue).

On dira que le langage $L_1 \subseteq X_1^*$ domine le langage $L_2 \subseteq X_2^*$, et nous noterons $L_1 > L_2$, s'il existe une transduction rationnelle τ telle que $L_2 = \tau(L_1)$. On dira que L_1 est rationnellement équivalent à L_2 si L_1 domine L_2 et L_2 domine L_1, ce que nous noterons $L_1 \equiv L_2$. Enfin on dira que L_1 et L_2 sont rationnellement incomparables si L_1 ne domine pas L_2 et L_2 ne domine pas L_1.

Définition 1.6

Un cône rationnel (fidèle) est une famille de langages fermée par transduction rationnelle (fidèle).

Du théorème de Nivat et du théorème 1.5 il résulte alors qu'une famille de langages $\mathcal{L}$ est un cône rationnel (fidèle) si et seulement si $\mathcal{L}$ est fermée par homomorphisme (non effaçant), homomorphisme inverse et intersection avec un langage rationnel [Lat1] et nous retrouvons la définition donnée dans [Gi2].

Définition 1.7

Pour toute famille de langages $\mathcal{L}$, $\mathcal{C}(\mathcal{L})$ (resp. $\mathcal{C}^f(\mathcal{L})$) désignera le plus petit cône rationnel (fidèle) contenant $\mathcal{L}$.

Un cône rationnel (fidèle) $\mathcal{L}$ est principal s'il existe un langage L tel que $\mathcal{L} = \mathcal{C}(\{L\})$ ($\mathcal{L} = \mathcal{C}^f(\{L\})$). Nous noterons alors abusivement $\mathcal{C}(L)$ (resp. $\mathcal{C}^f(L)$). Le langage L est appelé générateur (fidèle) de $\mathcal{L}$.

Un cône rationnel (fidèle) fermé par union est appelé un "full semi AFL" ("semi AFL") dans [GG1]. Nous noterons par $\mathcal{C}^U(\mathcal{L})$ (resp. $\mathcal{C}^{fU}(\mathcal{L})$) le plus petit cône rationnel (fidèle) fermé par union contenant $\mathcal{L}$.

On a alors le résultat classique :

Lemme 1.1 [GG1]

Un cône rationnel (fidèle) principal est fermé par union.

Définition 1.8

Une famille agréable de langages (FAL en abrégé, "full AFL" dans [GG1])(resp. une FAL fidèle, "AFL" dans [GG1]) est un cône rationnel (resp. fidèle) fermé par union, produit et étoile (resp. étoile propre).

$\mathfrak{F}(\mathcal{L})$ (resp $\mathfrak{F}^f(\mathcal{L})$) désignera la plus petite FAL (fidèle) contenant la famille de langages $\mathcal{L}$.

Une FAL (fidèle) $\mathcal{L}$ est dite principale s'il existe un langage L tel que $\mathcal{L} = \mathfrak{F}(\{L\})$ (resp. $\mathfrak{F}^f(L)$) que nous noterons $\mathfrak{F}(L)$ (resp. $\mathfrak{F}^f(L)$). Le langage L sera dit générateur de la FAL (resp. générateur fidèle) de la FAL (fidèle) $\mathcal{L}$.

Le lien entre les notions de cônes rationnels (fidèles) et de FAL (fidèles) est précisé par le théorème suivant :

Théorème 1.7 [GG2]

Etant donné un langage L sur X et un symbole $c \notin X$, on a :

- $\mathfrak{F}(L) = \mathcal{C}((Lc)^*)$
- $\mathfrak{F}^f(L) = \mathcal{C}^f((Lc)^*)$

Il en résulte que si $\mathcal{L}$ est un cône rationnel (fidèle) principal alors $\mathfrak{F}(\mathcal{L})$ (resp. $\mathfrak{F}^f(\mathcal{L})$) est une FAL (fidèle) principale. Cependant la réciproque est fausse en général.

Nous rappellerons enfin les deux résultats ci-dessous :

Lemme 1.2 [GG2]

La plus petite FAL (fidèle) engendrée par une famille finie de langages est principale.

Le plus petit cône rationnel (fidèle) engendré par une famille finie de langages est principal.

Théorème 1.8 [Gr1]

Une FAL (resp. FAL fidèle, cône rationnel fermé par union, cône rationnel fidèle fermé par union) est non principale si et seulement si il existe une suite infinie de FAL (resp. FAL fidèles, cônes rationnels, cônes rationnels fidèles) $\mathcal{L}_i$ strictement croissante telle que $\mathcal{L} = \bigcup_{i \geq 1} \mathcal{L}_i$.

Rappelons que la suite $\mathcal{L}_i$ est strictement croissante si et seulement si $\mathcal{L}_i \subsetneq \mathcal{L}_{i+1}$, $\forall i \geq 1$.

C) RAPPELS SUR LES OPERATIONS ET OPERATEURS SUBSTITUTION, CHEVRON ET CROCHET

1. SUBSTITUTION

Définition 1.9

Soit X un alphabet fini et pour tout $a \in X$ soit X_a un alphabet fini et $\sigma(a)$ un langage inclus dans X_a^*. Si $\sigma(\varepsilon) = \varepsilon$ et $\sigma(x_1x_2\ldots x_n) = \sigma(x_1)\ \sigma(x_2)\ldots\ \sigma(x_n)$ pour tout $x_1,\ldots,x_n \in X$, alors l'application σ de X_1^* dans $2^{(\bigcup_{a \in X} X_a)^*}$ est appelée une substitution sur X.

On étend de façon canonique σ sur 2^{X^*} par $\sigma(L) = \bigcup_{w \in L} \sigma(w)$.

La substitution σ est propre si $\varepsilon \notin \sigma(a)$ pour tout $a \in X$.

Pour toutes familles de langages $\mathscr{L}_1$ et $\mathscr{L}_2$ nous désignerons par $\underline{\mathscr{L}}_1\ \mathscr{L}_2$ (resp. $\underline{\mathscr{L}}_1^f\ \mathscr{L}_2$) la famille de langages.

$\underline{\mathscr{L}}_1\ \mathscr{L}_2 = \{\sigma(L) \mid L \in \mathscr{L}_1$, σ est une substitution sur X_L^* telle que $\sigma(a) \in \mathscr{L}_2$, pour tout $a \in X_L\}$.

(resp. $\underline{\mathscr{L}}_1^f\ \mathscr{L}_2 = \{\sigma(L) \mid L \in \mathscr{L}_1$, σ est une substitution propre sur X_L^* telle que $\sigma(a) \in \mathscr{L}_2$, pour tout $a \in X_L\}$).

Nous appellerons $\underline{\mathscr{L}}_1$ (resp. $\underline{\mathscr{L}}_1^f$) opérateur substitution (fidèle) dans $\mathscr{L}_1$.

On a alors les théorèmes fondamentaux

Théorème 1.9 [GS1]

1) Si $\mathscr{L}_1$ et $\mathscr{L}_2$ sont tous deux des cônes rationnels fidèles (resp.

fermés par union, FAL fidèles) alors $\underline{\mathcal{L}}_1^f\ \mathcal{L}_2$ et $\underline{\mathcal{L}}_1\ \mathcal{L}_2$ sont des cônes rationnels fidèles (resp. cônes rationnels fidèles fermés par union, FAL fidèles).

2) Si $\mathcal{L}_1$ est un cône rationnel fidèle et $\mathcal{L}_2$ un cône rationnel (resp. $\mathcal{L}_1$ cône rationnel fidèle fermé par union et $\mathcal{L}_2$ un cône rationnel fermé par union, $\mathcal{L}_1$ une FAL fidèle et $\mathcal{L}_2$ une FAL) alors $\underline{\mathcal{L}}_1\ \mathcal{L}_2$ est un cône rationnel (resp. cône rationnel fermé par union, FAL). De plus si $\mathcal{L}_1$ est fermé par homomorphisme quelconque alors

$$\underline{\mathcal{L}}_1^f\ \mathcal{L}_2 = \underline{\mathcal{L}}_1\ \mathcal{L}_2.$$

Théorème 1.10 [Bo1]

Le plus petit cône rationnel (fidèle) fermé par union et produit et contenant la famille de langages $\mathcal{L}$ est $\underline{\text{Fin}}\ \mathcal{C}(L)$ (resp. $\underline{\text{Fin}}_0\ \mathcal{C}^f(\mathcal{L})$).

Théorème 1.11 [GG1]

La plus petite FAL (fidèle) contenant la famille de langages $\mathcal{L}$ est $\mathcal{F}(\mathcal{L}) = \underline{\text{Rat}}\ \mathcal{C}(\mathcal{L})$ (resp. $\mathcal{F}^f(\mathcal{L}) = \underline{\text{Rat}}_0\ \mathcal{C}^f(\mathcal{L})$).

Plus généralement Fin (resp. Rat) est l'opérateur de fermeture de la famille $\mathcal{L}$ par union et produit (resp. et étoile).

Une substitution particulière à des propriétés remarquables qui seront utilisées dans la suite :

Définition 1.10

Une substitution σ est une "bonne substitution" sur X s'il existe un langage L_2 tel que $\sigma(a) = aL_2$ pour tout $a \in X$. Nous noterons alors pour tout langage $L_1 \subseteq X^*$:

$$\sigma(L_1) = s(L_1,L_2) = \{a_1x_1 \ldots a_nx_n \mid n \in \mathbb{N}, a_1a_2\ldots a_n \in L_1$$
$$x_i \in L_2,\ 1 \leq i \leq n\}.$$

On a alors :

Théorème 1.11 [Gr2] [Beau1] [Lat2] (Lemme syntaxique)

Soient L_1 et L_2 deux langages sur des alphabets disjoints et c un nouveau symbole. Soit $\mathcal{L}_2$ un cône rationnel et $\mathcal{L}_1$ une famille de langage fermée par Fin (i.e. Fin $\mathcal{L}_2 \subseteq \mathcal{L}_2$). Alors $s(L_1,L_2) \in \underline{\mathcal{L}}_1 \mathcal{L}_2$ implique soit $L_1 \in \mathcal{L}_1$ soit $(L_2c)^* \in \mathcal{L}_2$.

Il est a noter que si $\mathcal{L}_1$ et $\mathcal{L}_2$ sont deux cônes rationnels alors

$$\underline{\mathcal{L}}_1 \mathcal{L}_2 = \{s(L_1,L_2) \mid L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2\}.$$

D'autre part si $\mathcal{L}_1$ = Rat on en déduit :

Corollaire 1.1 [Gr2]

Si $\mathcal{L}$ est un cône rationnel, L_1 et L_2 deux langages sur des alphabets disjoints, alors $s(L_1,L_2) \in \underline{\text{Rat}}\ \mathcal{L}$ implique que L_1 ou L_2 appartient à $\mathcal{L}$.

On dira que $\mathcal{L}_2$ est fermé par substitution dans $\mathcal{L}_1$ si $\underline{\mathcal{L}}_1 \mathcal{L}_2 \subseteq \mathcal{L}_2$.

Il résulte alors du corollaire 1.1 qu'un cône rationnel $\mathcal{L}$ est clos par substitution si et seulement si $\mathcal{L} = \mathcal{F}(\mathcal{L})$. On en déduit :

Corollaire 1.2 [Gr3]

Si $\mathcal{L}$ est un cône rationnel qui n'est pas une FAL, $\mathcal{F}(\mathcal{L})$ n'est pas fermé par substitution.

Corollaire 1.3 [Gr2]

Si $\mathcal{L}$ est un cône rationnel non fermé par substitution, alors le plus petit cône rationnel contenant $\mathcal{L}$ fermé par substitution est non principal.

2. CHEVRON ET CROCHET

Définition 1.11 [BCN]

Pour tout langage L et tout élément $A \in Rat_2(X)$ nous appellerons crochet de L par A le langage

$$C(L,A) = \{f_1 \ell \tilde{f}_2 | \ell \in L, (f_1,f_2) \in Rat_2(X)\}$$

où $\tilde{f}_2$ désigne l'image miroir de f_2.

Si a et b sont deux lettres n'appartenant pas à X, nous appellerons chevron du langage $L \subseteq X^*$, le langage

$$Ch(L) = \{a^n \ell b^n | n \in \mathbb{N}, \ell \in L\}$$

Ch(L) est donc défini à un isomorphisme alphabétique près.

Une famille de langages $\mathcal{L}$ sera dite fermée par crochet (resp. chevron) si pour tout langage $L \in \mathcal{L}$ et tout $A \in Rat_2(X)$, $C(L,A) \in \mathcal{L}$ (resp. $Ch(L) \in \mathcal{L}$).

Etant donné une famille de langage $\mathcal{L}$ nous noterons $C(\mathcal{L})$ (resp. $C^f(\mathcal{L})$) le plus petit cône rationnel (fidèle) contenant la famille $\mathcal{L}'$ des crochets de $\mathcal{L}$, i.e.

$$\mathcal{L}' = \{C(L,A) | L \in \mathcal{L}, A \in Rat_2(X)\}.$$

Nous noterons $Ch(\mathcal{L})$ (resp. $Ch^f(\mathcal{L})$) le plus petit cône rationnel (fidèle) contenant la famille $\mathcal{L}''$ des chevrons de $\mathcal{L}$, i.e.

$$\mathcal{L}'' = \{Ch(L) | L \in \mathcal{L}\}.$$

On a alors les résultats fondamentaux suivants :

Lemme 1.2 [BONi1]

Etant donné un langage $L \subseteq X^*$ et deux symboles a et b n'appartenant pas à X, si $\mathcal{L}$ est un cône rationnel (fidèle) fermé par union, la condition $Ch(L) = \{a^n \ell b^n | \ell \in L, n \in \mathbb{N}\} \in \underline{Rat}\ \mathcal{L}$ implique $L \in \mathcal{L}$ (resp. $Ch(L) \in \underline{Rat}_0\ \mathcal{L} \Rightarrow L \in \mathcal{L}$).

Il en résulte alors

Corollaire 1.4 [BONi1]

Si $\mathcal{L} = \mathcal{F}(\mathcal{S})$ (resp. $\mathcal{L} = \mathcal{F}^f(\mathcal{S})$) est une FAL (resp. FAL fidèle) fermée par chevron, alors

$$\mathcal{L} = \mathcal{C}^U(\mathcal{S}) \text{ (resp. } \mathcal{L} = \mathcal{C}^{fU}(\mathcal{S})).$$

Dans le cas du crochet, on a :

Proposition 1.1 [BCN]

Si $\mathcal{L}$ est un cône rationnel, la famille $C(\mathcal{L})$ est identique à la fermeture par union de la famille

$$\mathcal{L}' = \{C(L,A) | L \in \mathcal{L}, A \in Rat_2(X)\}.$$

Proposition 1.2 [BCN]

Soient L_1 et L_2 deux langages sur des alphabets disjoints et $\mathcal{L}$ un cône rationnel fermé par union. La condition $L_1L_2 \in C(\mathcal{L})$ implique que L_1 est rationnel ou $L_2 \in \mathcal{L}$.

Du lemme 1.2 et de la proposition 1.2, il résulte alors

Théorème 1.12 [BCN]

Soit $\mathcal{L}$ un cône rationnel fermé par union.

Les conditions :

a) $\mathcal{L} \neq Ch(\mathcal{L})$
b) $\mathcal{L} \neq C(\mathcal{L})$

impliquent respectivement :

a') Le plus petit cône rationnel contenant $\mathcal{L}$ fermé par union, produit et chevron est non principal. La plus petite FAL contenant $\mathcal{L}$ fermée par chevron est non principale.

b') Le plus petit cône rationnel contenant $\mathcal{L}$ fermé par union, produit et crochet est non principal. La plus petite FAL contenant $\mathcal{L}$ fermée par crochet est non principale.

Remarques

1) L'opérateur crochet qui associe la famille $C(\mathcal{L})$ à toute famille de langages $\mathcal{L}$ est idempotent, i.e. $C(C(\mathcal{L})) = C(\mathcal{L})$.

2) Nous verrons ultérieurement que l'on peut définir des opérateurs crochet fidèle et chevron fidèle qui ont des propriétés analogues à celles des opérateurs crochet et chevron. Nous verrons enfin qu'on peut établir pour l'opération chevron un lemme "syntaxique" analogue au théorème 1.11 pour les bonnes substitutions.

CHAPITRE 2

OPÉRATIONS ET OPÉRATEURS

A) INTRODUCTION

B) DEFINITIONS PRELIMINAIRES, PROPRIETES ELEMENTAIRES

C) OPERATEURS REGULIERS

D) OPERATIONS INDEPENDANTES D'UN OPERATEUR

E) OPERATIONS INDEPENDANTES DE Rat

F) OPERATIONS DE DEGRE $p \geq 0$

G) OPERATIONS SYNTAXIQUES

H) FAL DECOMPOSABLES

A) INTRODUCTION

La théorie algébrique des langages a utilisé, tout d'abord, des opérations sur les langages et des opérateurs sur les familles de langages extrêmement généraux : transductions rationnelles, fermeture rationnelle d'un ensemble, cônes rationnels ou FAL engendrés. L'étude systématique de familles de langages classiques, initialement définies par grammaires ou automates, a nécessité, peu à peu, l'introduction d'opérations et d'opérateurs plus spécifiques (substitution, insertion, crochet, chevron...). Dans tous les cas des similitudes de propriétés de ces opérations et de ces opérateurs sont apparues à des niveaux les plus divers. Ainsi une caractéristique importante de certaines de ces opérations est la suivante : si $\mathcal{L}$ est un cône rationnel non fermé pour une opération t alors la fermeture de $\mathcal{L}$ par t est non principale. A partir de cette remarque, S. Greibach a défini la classe des opérations syntaxiques binaires qui vérifient de telles propriétés (substitution, insertion,...) [Gr3]. Or, comme le soulignent Boasson, Crestin et Nivat dans [BCN], les opérations chevron et crochet ont des propriétés analogues à celles des opérations binaires syntaxiques de [Gr3], mais ces propriétés ne peuvent être déduites de la théorie de S. Greibach. En effet d'une part le chevron et le crochet sont des opérations unaires et, d'autre part, le crochet est associé à un opérateur idempotent, i.e. $C(C(\mathcal{L})) = C(\mathcal{L})$.

Un des buts de notre thèse est de dégager les concepts fondamentaux qui déterminent ces similitudes de propriétés. Ceci implique donc la considération d'opérations d'arité n quelconque, $n \geq 1$, et d'opérations qui ne vérifient pas nécessairement des propriétés aussi fortes que celles des opérations syntaxiques de S. Greibach.

Plus précisément nous caractérisons tout d'abord les opérateurs T tels que si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels principaux de générateurs respectifs $L_1,\ldots,L_n$, alors $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est un cône rationnel principal dont un générateur est une fonction des langages $L_1,\ldots,L_n$. Ces opérateurs sont appelés opérateurs réguliers. Ainsi Rat -l'opérateur de fermeture rationnelle-, la substitution, l'insertion, le crochet et le chevron sont réguliers. Dans le cas de l'opérateur substitution ceci nous permet de retrouver

simplement des résultats de [GG2] et de généraliser un résultat de Beauquier [Beau1].

Dans une deuxième partie nous définissons les opérations indépendantes d'un opérateur : une opération n-aire t est indépendante de l'opérateur m-aire T' si pour tous langages $L_1,\ldots,L_n$ sur des alphabets disjoints deux à deux et tous cônes rationnels fermés par union $\mathscr{L}_1,\ldots,\mathscr{L}_m$, la condition $t(L_1,\ldots,L_n) \in T'(\mathscr{L}_1,\ldots,\mathscr{L}_m)$ implique qu'il existe $1 \leq i \leq n$ et $1 \leq j \leq m$ tels que $L_i \in \mathscr{L}_j$.

Ainsi le produit de deux langages est indépendant du crochet [BCN], le chevron, la substitution et l'insertion sont indépendantes de Rat [BoNi1,Gr3]. Ceci nous permet d'établir un théorème qui implique la non principalité de certains cônes rationnels. Nous retrouvons ainsi aussi bien les résultats de [Gr3] pour les opérations binaires syntaxiques que ceux de [BCN] pour le crochet et le chevron.

Deux cas particuliers se révèlent alors d'un grand intérêt. Le premier est celui des opérations indépendantes de Rat, l'opérateur de fermeture rationnelle. Nous établissons tout d'abord une condition nécessaire et suffisante pour que certaines FAL (fidèles) soient principales ce qui nous permet de simplifier les preuves de résultats connus et d'en démontrer de nouveaux. Nous montrons ensuite que la famille des langages rationnels est la seule FAL fermée par une opération indépendante de Rat et contenue dans la FAL engendrée par les langages bornés. On établit une propriété analogue pour les langages linéaires et, dans le chapitre suivant, nous verrons qu'il en est de même pour la FAL engendrée par $L_1 = \{wcw | w \in \{a,b\}^*\}$.

Le second cas particulier que nous étudions est celui des opérations indépendantes de l'opérateur auquel elle est associée. Ces opérations sont dites de degré 0 et comportent notamment le chevron et la substitution. Suivant la démarche de S. Greibach nous introduisons alors les opérations de degré $p > 0$ afin de traiter le cas de l'insertion. Nous établissons alors de nouveaux lemmes impliquant la non principalité de cônes rationnels.

Ce chapitre s'achève par l'étude des opérations de degré $p \geq 0$ et

indépendantes de Rat que nous appelons opérations syntaxiques. Les résultats de [Gr3] se généralisent alors simplement. Considérant enfin les FAL (fidèles) fermées par une opération syntaxique nous apportons la solution à un problème ouvert de S. Greibach dans [Gr3] : toute FAL (cône rationnel) fidèle principal fermée par une opération syntaxique n-aire, $n \geq 1$, contient une FAL (cône rationnel) fidèle maximale. Ceci nous permet, par exemple, de montrer que les familles $\mathcal{L}_{BNP}$ [BNP] et $\mathcal{L}_{DUP}$ [BG] contiennent des cônes rationnels fidèles maximaux.

En conclusion les trois concepts fondamentaux que nous dégageons dans ce chapitre (opérations régulières , opérations indépendantes d'un opérateur et opérations de degré $p \geq 0$) nous permettent d'expliquer à la fois la similitude de propriétés des diverses opérations considérées à ce jour en théorie des langages ainsi que d'établir un certain nombre de résultats nouveaux sur des familles de langages classiques.

B) DEFINITIONS PRELIMINAIRES, PROPRIETES ELEMENTAIRES

Définition 2.1

Soit $\hat{\sum}$ la famille de tous les langages définis sur les sous-ensembles finis de l'alphabet universel $\sum$.

Une opération n-aire t ($n \geq 1$) est une application de $\hat{\sum}^n$ dans $\hat{\sum}$, notée $t(L_1,\ldots,L_n)$ pour $L_1,\ldots,L_n$ appartenant à $\hat{\sum}$.

Une opération n-aire ($n \geq 1$) t est régulière si pour tous langages L_i et L_i' tels que L_i domine rationnellement L_i', $1 \leq i \leq n$, alors le langage $t(L_1,\ldots,L_n)$ domine rationnellement $t(L_1',\ldots,L_n')$, i.e. $L_i > L_i'$, $1 \leq i \leq n \Longrightarrow t(L_1,\ldots,L_n) > t(L_1',\ldots,L_n')$.

Définition 2.2

Un opérateur n-aire T ($n \geq 1$) est une application de $2^{\hat{\sum}^n}$ dans $2^{\hat{\sum}}$ qui à toutes familles de langages $\mathscr{L}_1,\ldots,\mathscr{L}_n$ fait correspondre une famille de langages notée $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$.

Un opérateur n-aire T est simple si pour toutes familles de langages $\mathscr{L}_1,\ldots,\mathscr{L}_n$ on a $\mathscr{L}_1 \cup \mathscr{L}_2 \cup \ldots \cup \mathscr{L}_n \subseteq T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$.

Un opérateur n-aire T est conique (fidèle) si pour tous cônes rationnels (fidèles) $\mathscr{L}_1,\ldots,\mathscr{L}_n$ la famille $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ forme un cône rationnel (fidèle) fermé par union.

Un opérateur conique (fidèle) n-aire T est associé à une opération n-aire t s'il est simple et si pour tous cône rationnels (fidèles) $\mathscr{L}_1, \mathscr{L}_2,\ldots,\mathscr{L}_n$ on a :

$$T(\mathscr{L}_1,\ldots,\mathscr{L}_n) = \mathscr{C}^{U}\{t(L_1,\ldots,L_n) \mid L_i \in \mathscr{L}_i,\ 1 \leq i \leq n\}$$

$$(T(\mathscr{L}_1,\ldots,\mathscr{L}_n) = \mathscr{C}^{fU}\{t(L_1,\ldots,L_n) \mid L_i \in \mathscr{L}_i,\ 1 \leq i \leq n\}.$$

Un opérateur T est régulier s'il est associé à une opération régulière.

Si $\mathscr{L}$ est une famille de langages, nous dirons que $\mathscr{L}$ est fermée pour une opération n-aire t, si pour tous langages $L_1,\ldots,L_n$ de $\mathscr{L}$ on a $t(L_1,\ldots,L_n) \in \mathscr{L}$.

On dira que $\mathscr{L}$ est fermée pour l'opérateur n-aire T si $T(\mathscr{L},\ldots,\mathscr{L}) \subseteq \mathscr{L}$.

Notations

Si t (resp. T) est une opération (resp. un opérateur) n-aire nous noterons t(L) (resp. $T(\mathscr{L})$) le langage $t(L_1,L_2,\ldots,L_n)$ avec $L_1 = L_2 = \ldots = L_n = L$ (resp. la famille $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ avec $\mathscr{L}_1 = \ldots = \mathscr{L}_n = \mathscr{L}$).

Dans la suite nous utiliserons souvent un même symbole pour désigner une opération et son opérateur associé. L'utilisation de lettres majuscules droites pour les langages et rondes pour les familles de langages lèvera toute ambiguité.

Remarque

Il est clair que l'on peut associer un opérateur conique (fidèle) T à toute opération n-aire en posant

$$T(\mathscr{L}_1,\ldots,\mathscr{L}_n) = \mathscr{C}^{U}\{t(L_1,\ldots,L_n) \mid L_i \in \mathscr{L}_i,\ 1 \le i \le n\}$$

(resp. $T(\mathscr{L}_1,\ldots,\mathscr{L}_n) = \mathscr{C}^{fU}\{t(L_1,\ldots,L_n) \mid L_i \in \mathscr{L},\ 1 \le i \le n\}$.

Toutefois l'opérateur obtenu n'est pas nécessairement simple. Il suffit de considérer l'opération $\cap$ d'intersection de deux langages. Alors l'opérateur binaire T défini par

$T(\mathscr{L}_1,\mathscr{L}_2) = \mathscr{C}^{U}(L_1 \cap L_2 \mid L_1 \in \mathscr{L}_1, L_2 \in \mathscr{L}_2)$ n'est pas simple.

Exemples

1. Les opérateurs unaires Fin et Rat (resp. Fin_0 et Rat_0) qui à toute famille de langages $\mathcal{L}$ associent respectivement sa fermeture par union et produit et sa fermeture par union, produit et étoile (propre) sont des opérateurs coniques (fidèles) simples. D'après le théorème 1.7 on a pour tout cône rationnel (fidèle $\mathcal{L}$),

$$\text{Rat}\ \mathcal{L} = \mathfrak{C}^{U}\{(Lc)^* | L \in \mathcal{L}\}$$

(resp. $\text{Rat}_0\ \mathcal{L} = \mathfrak{C}^{fU}\{(Lc)^* | L \in \mathcal{L}\}$), l'opérateur Rat est donc associé à l'opération étoile "marquée" qui à tout langage L associe le langage $(Lc)^*$ pour un nouveau symbole c. Par contre il n'existe aucune opération unaire t telle que pour tout cône rationnel $\mathcal{L}$ on ait

$$\text{Fin}\ \mathcal{L} = \mathfrak{C}^{U}\{t(L) | L \in \mathcal{L}\}.$$

Ceci prouve qu'en général la réciproque de la remarque précédente est fausse.

Notons enfin que l'opération étoile marquée est trivialement régulière.

2. L'opérateur unaire chevron est associé à l'opération du même nom. Cet opérateur est simple, conique et régulier. En effet si $L_1 > L_2$, alors il est facile de construire une transduction rationnelle τ telle que $\tau(Ch(L_1)) = Ch(L_2)$.

3. L'opérateur unaire crochet est simple mais il n'est pas associé à "l'opération" crochet de L par $A \in \text{Rat}_2$ qui n'est pas une opération au sens où nous l'avons défini. Nous verrons par la suite que cet opérateur peut être associé à l'opération C(L), crochet du langage $L \subseteq X^*$, définie par :

Si a et b n'appartiennent pas à X alors :

$$C(L) = \{w\ell\tilde{w} | w \in \{a,b\}^*, \ell \in L\} \qquad [\text{Rod } 4]$$

Il est facile de vérifier que cette dernière opération est

régulière.

4. L'opérateur substitution de $\mathcal{L}_2$ dans $\mathcal{L}_1$ est un opérateur binaire simple conique (fidèle) qui est associé à l'opération substitution, mais également à l'opération de "bonne substitution". En fait à tout cône rationnel (fidèle) $\mathcal{L}$ on peut associer l'opérateur unaire $\underline{\mathcal{L}}$ de substitution dans $\mathcal{L}$ qui est simple mais n'est pas associé à une opération unaire dans le cas général (cf Fin dans l'exemple 1).

Notations

Pour toute opération t (resp. opérateur T) n-aire et toute famille de langages $\mathcal{L}$, $\mathcal{C}_t^{fU}(\mathcal{L})$ (resp. $\mathcal{C}_T^{fU}(\mathcal{L})$) désignera le plus petit cône rationnel fidèle contenant $\mathcal{L}$ et fermé par union et t (resp. T), $\mathcal{C}_t^{U}(\mathcal{L})$ (resp. $\mathcal{C}_T^{U}(\mathcal{L})$) désignera le plus petit cône rationnel contenant $\mathcal{L}$ fermé par union et t (resp. T), $\mathcal{F}_t^{f}(\mathcal{L})$ (resp. $\mathcal{F}_T^{f}(\mathcal{L})$) la plus petite FAL fidèle contenant $\mathcal{L}$ fermée par t (resp. T) et $\mathcal{F}_t(\mathcal{L})$ (resp. $\mathcal{F}_T(\mathcal{L})$) la plus petite FAL contenant $\mathcal{L}$ et fermée par t (resp. T).

Définition 2.3

Pour tout opérateur n-aire ($n \geq 1$) T nous désignerons, pour tout $p \geq 1$, par T^p, φ_T^p (resp. φ_T^{fp}) et ψ_T^p (resp. ψ_T^{fp}) les opérateurs définis par :

Si $\mathcal{L}_1, \ldots, \mathcal{L}_n$ sont des familles de langages alors

$$T^1(\mathcal{L}_1, \ldots, \mathcal{L}_n) = T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$$

$$\varphi_T^1(\mathcal{L}_1, \ldots, \mathcal{L}_n) = T(\underline{\text{Fin}}\,\mathcal{L}_1, \ldots, \underline{\text{Fin}}\,\mathcal{L}_n)$$

$$(\text{resp. } \varphi_T^{f1}(\mathcal{L}_1, \ldots, \mathcal{L}_n) = T(\underline{\text{Fin}_0}\,\mathcal{L}_1, \ldots, \underline{\text{Fin}_0}\,\mathcal{L}_n))$$

$$\psi_T^1(\mathcal{L}_1, \ldots, \mathcal{L}_n) = T(\underline{\text{Rat}}\,\mathcal{L}_1, \ldots, \underline{\text{Rat}}\,\mathcal{L}_n)$$

$$(\text{resp. } \psi_T^{f1}(\mathcal{L}_1, \ldots, \mathcal{L}_n) = T(\underline{\text{Rat}_0}\,\mathcal{L}_1, \ldots, \underline{\text{Rat}_0}\,\mathcal{L}_n)$$

et pour $p \geq 1$:

$$T^{p+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n) = T(T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n))$$

$$\varphi_T^{p+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n) = T(\underline{\text{Fin}}\ \varphi_T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n))$$

(resp. $\varphi_T^{fp+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n) = T(\underline{\text{Fin}}_0\ \varphi_T^{fp}(\mathcal{L}_1,\ldots,\mathcal{L}_n))$

$$\psi_T^{p+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n) = T(\underline{\text{Rat}}\ \psi_T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n))$$

(resp. $\psi_T^{fp+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n) = T(\underline{\text{Rat}}_0\ \psi_T^{fp}(\mathcal{L}_1,\ldots,\mathcal{L}_n)$

Définition 2.4

Pour toutes familles de langages $\mathcal{L}_1$ et $\mathcal{L}_2$ soit

$$\mathcal{L}_1 \vee \mathcal{L}_2 = \{L_1 \cup L_2 | L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2\}.$$

On a alors :

Lemme 2.1

Soit T un opérateur conique (fidèle) simple et $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnel (fidèle). Alors

$$\mathcal{C}_T^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n)$$

(resp. $\mathcal{C}_T^{fU}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n)$)

$$\mathcal{C}_T^{U\bullet}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} \varphi_T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n)$$

(resp. $\mathcal{C}_T^{fU\bullet}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} \varphi_T^{fp}(\mathcal{L}_1,\ldots,\mathcal{L}_n)$)

$$\mathcal{F}_T(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} \psi_T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n)$$

(resp. $\mathcal{F}_T^f(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) = \bigcup_{p \geq 1} \psi_T^{fp}(\mathcal{L}_1,\ldots,\mathcal{L}_n)$).

Démonstration

Nous ne la ferons que pour $\mathcal{C}_T^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$, elle est identique dans les autres cas. Il est clair que $\mathcal{C} = \bigcup_{p \geq 1} T^p(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est inclus dans $\mathcal{C}_T^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$. Réciproquement $\mathcal{C}$ est un cône rationnel fermé par union et pour tout langage L appartenant à $\mathcal{C}_T^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$ il existe un entier p tel que L appartient à $T^p(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) \subseteq T^{p+1}(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ puisque T est simple. Donc L appartient à $\mathcal{C}$.

■

Nous terminerons cette première partie par le lemme :

Lemme 2.2

Soient T_1 et T_2 deux opérateurs simples n-aires. Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des familles de langages telles que $T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ est fermé par T_2 et $T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ est fermé par T_1, alors

$$T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n)) = T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)).$$

Démonstration

Puisque T_2 est simple, $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n \subseteq T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ et donc $T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n) \subseteq T_1(\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n) \subseteq T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n))$. Comme

$T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ est fermé pour T_2 il en résulte que

$$T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)) \subseteq T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n))$$

La réciproque est identique.

■

Il en résulte alors par exemple :

Corollaire 2.1 [Rod1]

Pour toute famille de langages recopiable $\mathcal{L}$

1) $\underline{\text{Fin}}\ \mathcal{E}(\mathcal{L}) = \mathcal{E}(\underline{\text{Fin}}(\mathcal{L}))$.

2) $\underline{\text{Rat}}\ \mathcal{E}(\mathcal{L}) = \mathcal{E}(\underline{\text{Rat}}(\mathcal{L}))$.

C) OPERATEURS REGULIERS

Un opérateur n-aire T est régulier s'il est associé à une opération régulière t, i.e. une opération telle que si $L_1,\ldots,L_n$ sont des langages qui dominent rationnellement les langages $L'_1,\ldots,L'_n$ alors $t(L_1,\ldots,L_n)$ domine rationnellement $t(L'_1,\ldots,L'_n)$. Le résultat fondamental de cette section affirme que si $\mathscr{L}_1,\ldots,\mathscr{L}_n$ sont des cônes rationnels principaux et T un opérateur n-aire régulier alors $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ est un cône rationnel principal (théorème 2.1). De plus si $L_1,\ldots,L_n$ sont des générateurs respectifs de $\mathscr{L}_1,\ldots,\mathscr{L}_n$ alors $t(L_1,\ldots,L_n)$ est un générateur de $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$.

Nous avons vu dans la partie B que Rat, le chevron et le crochet sont des opérateurs réguliers. Le théorème 2.1 nous permet donc de retrouver simplement des résultats classiques de [Gi2], [BCN], [BoNi1] ainsi que des résultats de S. Greibach [Gr3] pour l'insertion ou "l'homomorphic replication".

Par contre l'opération "bonne substitution" n'est pas régulière. Nous montrons alors que l'opérateur substitution peut être associé à une opération régulière notée t_s et définie par $t_s(L_1,L_2) = s(L_1,(L_2c)^+)$ où s désigne l'opération de bonne substitution. Ceci nous permet de retrouver simplement un résultat de [GG2] et de déduire directement du lemme syntaxique de S. Greibach [Gr2] une version plus générale du lemme syntaxique de Beauquier [Beau1] cependant moins forte que celle de Latteux [Lat2]. De façon générale nous montrons enfin qu'un opérateur T est régulier si et seulement si pour tous cônes rationnels principaux $\mathscr{L}_1,\ldots,\mathscr{L}_n$ de générateurs respectifs $L_1,\ldots,L_n$, le cône rationnel $T(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ est principal de générateur $t'(L_1,\ldots,L_n)$ pour quelque opération t'. En fait T est associé à l'opération t' qui est régulière (théorème 2.2).

Théorème 2.1

Soit T un opérateur n-aire (fidèle) régulier simple associé à l'opération n-aire t. Soient $L_1,L_2,\ldots,L_n$ des langages, alors

$$T(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n)) = \mathcal{C}(t(L_1,\ldots,L_n))$$

(resp. $T(\mathcal{C}^f(L_1),\ldots,\mathcal{C}^f(L_n)) = \mathcal{C}^f(t(L_1,\ldots,L_n))$).

Démonstration

On a évidemment $t(L_1,\ldots,L_n) \in T(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n))$ et comme T est conique $\mathcal{C}(t(L_1,\ldots,L_n)) \subseteq T(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n))$.

Réciproquement si $L \in T(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n))$ il existe un entier p, np langages L_{ij}, $1 \le i \le n$, $1 \le j \le p$, et p transductions rationnelles τ_j, $1 \le j \le p$ telles que

$$L = \bigcup_{j=1}^{p} \tau_j\, t(L_{1j},\ldots,L_{nj})$$

avec $L_{ij} \in \mathcal{C}(L_i)$, $1 \le i \le n$, $1 \le j \le p$.

Pour tout j, il existe donc n transductions rationnelles τ_{ij}, $1 \le i \le n$, telles que $L_{ij} = \tau_{ij}(L_i)$. Comme t est une opération régulière, il en résulte que $t(L_1,\ldots,L_n)$ domine rationnellement $t(L_{1j},\ldots,L_{nj})$ pour tout j et par conséquent que le langage $\tau_j\, t(L_{1j},\ldots,L_{nj})$ appartient au cône rationnel $\mathcal{C}(t(L_1,\ldots,L_n))$ qui est fermé par union et donc contient L.

La preuve est identique dans le cas d'un opérateur fidèle.

■

On en déduit immédiatement les corollaires :

Corollaire 2.2

Soient $\mathcal{L}_1,\ldots,\mathcal{L}_n$, n cônes rationnels (fidèles) principaux de générateurs respectifs $L_1,\ldots,L_n$ et T un opérateur (fidèle) régulier n-aire simple associé à l'opération t. Alors $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est un cône rationnel (fidèle) de générateur $t(L_1,\ldots,L_n)$.

Corollaire 2.3

Un cône rationnel (fidèle) principal $\mathcal{L}$ est fermé pour une opération régulière t si et seulement si tout générateur L de $\mathcal{L}$ est rationnellement équivalent à t(L).

Corollaire 2.4

Soit T un opérateur (fidèle) n-aire régulier et simple et $\mathcal{L}$ une famille de langages, alors $T(\mathcal{C}(\mathcal{L})) = \mathcal{C}^{U}(T(\mathcal{L}))$ (resp. $T(\mathcal{C}^{f}(\mathcal{L})) = \mathcal{C}^{fU}(T(\mathcal{L}))$).

Exemples

1) Nous avons vu que le chevron est un opérateur régulier. Il en résulte que si $\mathcal{L}$ est un cône rationnel principal de générateur L alors Ch(L) est un générateur du cône rationnel $Ch(\mathcal{L})$. De même comme l'opération crochet est régulière, $C(L) = \{w\ell\tilde{w} \mid w \in \{a,b\}^*, \ell \in L\}$ est un générateur de $C(\mathcal{L})$.

2) Par contre l'opération de bonne substitution n'est pas régulière. En effet le langage {a} domine rationnellement le langage {ab} mais $s(\{a\}, S_2) = aS_2$ avec $S_2 = \{wc\tilde{w} \mid w \in \{a,b\}^*\}$ ne domine pas $s(\{ab\}, S_2) = aS_2bS_2$.

Toutefois on peut associer à $\underline{\mathcal{L}_1\,\mathcal{L}_2}$ une opération qui est régulière :

Lemme 2.3

Soient $\mathcal{L}_1$ et $\mathcal{L}_2$ deux cônes rationnels (fidèles) alors

$$\underline{\mathcal{L}_1\,\mathcal{L}_2} = \{s(L_1,(L_2c)^+) \mid L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2,\ c \notin X_{L_2}\},$$

i.e. l'opérateur $\underline{\mathcal{L}_1\,\mathcal{L}_2}$ est associé à l'opération $t_s(L_1,L_2) = s(L_1,(L_2c)^+)$.

Démonstration

Comme $\mathcal{L}_1$ et $\mathcal{L}_2$ sont des cônes rationnels (fidèles) on peut supposer que L_1 et L_2 sont sur des alphabets disjoints. Il est clair alors que si $s(L_1,(L_2c)^+) \in \underline{\mathcal{L}}_1\,\mathcal{L}_2$ alors $s(L_1,L_2) \in \underline{\mathcal{L}}_1\,\mathcal{L}_2$. Réciproquement montrons que si $s(L_1,L_2) \in \underline{\mathcal{L}}_1\,\mathcal{L}_2$ alors $s(L_1,(L_2c)^+) \in \underline{\mathcal{L}}_1\,\mathcal{L}_2$ pour $c \notin X_{L_2}$ et $X_{L_1} \cap X_{L_2} = \emptyset$. Comme $\mathcal{L}_1$ est un cône rationnel, si d est un nouveau symbole, le langage $L_1' = \{a_1 d^{n_1} c\ a_2 d^{n_2} c \ldots a_p d^{n_p} c \mid n_i > 0,\ 1 \le i \le p,\ a_1 a_2 \ldots a_p \in L_1\}$ appartient à $\mathcal{L}_1$. Soit alors σ la substitution définie par $\sigma(a) = \{a\}$ $\forall a \in X_{L_1}$, $\sigma(d) = L_2$ et $\sigma(c) = \{c\}$. Il est alors clair que $\sigma(L_1')$ est égal à $s(L_1,(L_2c)^+)$ et comme $\sigma(L_1') \in \underline{\mathcal{L}}_1\,\mathcal{L}_2$, le résultat est établi.

∎

On a alors :

Lemme 2.4

L'opération binaire t_s définie par $t_s(L_1,L_2) = s(L_1,(L_2c)^+)$ est régulière.

Démonstration

Il suffit de montrer que si $L_1 > L_1'$ alors $t_s(L_1,L_2) > t_s(L_1',L_2)$, la preuve de $t_s(L_1,L_2) > t_s(L_1',L_2')$ si $L_1 > L_1'$ et $L_2 > L_2'$ en découlant trivialement. On peut supposer, sans nuire à la généralité que L_1 et L_2 sont sur des alphabets disjoints puisque si M_2 est une recopie de L_2 alors $\underline{\mathcal{C}(L_1)}\ \mathcal{C}(L_2)$ est égal à $\underline{\mathcal{C}(L_1)}\ \mathcal{C}(M_2)$. Si L_1 domine rationnellement L_1', alors il existe, d'après le théorème de Nivat, deux homomorphismes alphabétiques φ et ψ et un langage régulier $R \subseteq Z^*$ tel que :

$$L_1' = \psi(\varphi^{-1}(L_1) \cap R).$$

Soient alors φ_1 et ψ_1 les homomorphismes définis par :

$$\varphi_1(x) = \varphi(x), \quad \forall x \in Z$$

$$\varphi_1(a) = a \quad , \forall a \in X_{L_2}$$

$$\psi_1(x) = \psi(x), \quad \forall x \in Z$$

$$\psi_1(a) = a \quad , \forall a \in X_{L_2}$$

et soit $R_1 = \theta^{-1}(R) \cap (Z\, X_{L_2}^*\, c)^*$ avec

$$\theta(x) = x \qquad \forall x \in Z$$

$$\theta(a) = \varepsilon \qquad \forall a \notin X_L \cup \{c\}.$$

On vérifie alors sans difficulté que

$$s(L_1', (L_2c)^+) = \psi_1(\varphi_1^{-1}(s(L_1, (L_2c)^+)) \cap R_1),$$

ce qui prouve le lemme.

■

On retrouve ainsi simplement le résultat de [GG2] :

Corollaire 2.5 [GG2]

Soient L_1 et L_2 deux langages non vides sur des alphabets disjoints, avec $L_1 \neq \{\varepsilon\}$. Si c est un nouveau symbole

$$\underline{\mathcal{C}^f(L_1)}\ \mathcal{C}^f(L_2) = \mathcal{C}^f(s(L_1,(L_2c)^+))$$

et

$$\underline{\mathcal{C}(L_1)}\ \mathcal{C}(L_2) = \mathcal{C}(s(L_1,(L_2c)^+)).$$

Le lemme 2.3 permet également de retrouver une version un peu plus faible du lemme syntaxique de M. Latteux [Lat2] à partir du lemme syntaxique de Beauquier [Beau1].

Corollaire 2.6

Si $\mathcal{L}_1$ et $\mathcal{L}_2$ sont deux cônes rationnels, L_1 et L_2 deux langages sur des alphabets disjoints et c un nouveau symbole, la condition $s(L_1,L_2) \in \underline{\mathcal{L}_1}\ \mathcal{L}_2$ implique que $L_1 \in \mathcal{L}_1$ ou $(L_2c)^+ \in \mathcal{L}_2$.

De façon générale nous avons le résultat fondamental suivant :

Théorème 2.1 bis

Un opérateur n-aire T est régulier si et seulement si il existe une opération t' telle que pour tous cônes rationnels principaux $\mathcal{L}_1,\ldots,\mathcal{L}_n$ de générateurs respectifs $L_1,\ldots,L_n$, le cône rationnel $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est principal de générateur $t'(L_1,\ldots,L_n)$.

Ce théorème résulte directement du lemme suivant :

Lemme 2.6

Soit T un opérateur conique et simple n-aire et t' une opération telle que pour tous cônes rationnels principaux $\mathcal{L}_1,\ldots,\mathcal{L}_n$ de générateurs respectifs $L_1,\ldots,L_n$, le cône rationnel $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est principal de générateur $t'(L_1,\ldots,L_n)$. Alors T est associé à l'opération t' et t' est régulière.

Démonstration

D'après nos hypothèses on a trivialement

$$T(\mathcal{N}_1,\ldots,\mathcal{N}_n) = \mathcal{C}^U(\{t'(M_1,\ldots,M_n) \mid M_i \in \mathcal{N}_i,\ 1 \leq i \leq n\})$$

pour tous cônes rationnels $\mathcal{N}_1,\ldots,\mathcal{N}_n$, donc T est associé à t'.

Soient alors n langages $L_1,\ldots,L_n$ qui dominent rationnellement les langages $L'_1,\ldots,L'_n$. Pour tout i, $1 \leq i \leq n$ on a donc $\mathcal{C}(L'_i) \subseteq \mathcal{C}(L_i)$ et d'après nos hypothèses $T(\mathcal{C}(L'_1),\ldots,\mathcal{C}(L'_n)) \subseteq T(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n))$ i.e. $t'(L'_1,\ldots,L'_n)$ est dominé rationnellement par $t'(L_1,\ldots,L_n)$ et donc t' est régulière.

■

D) OPERATION INDEPENDANTE D'UN OPERATEUR

Nous avons vu au chapitre 1 que du lemme syntaxique pour la substitution (théorème 1.11 et corollaire 1.1) on peut déduire (corollaire 1.3) que si $\mathcal{L}$ est un cône rationnel non fermé par substitution, alors le plus petit cône rationnel contenant $\mathcal{L}$ et fermé par substitution est non principal. Dans [Gr3] S. Greibach donne une formalisation de ce processus pour les opérations binaires qui vérifient des propriétés analogues à celles du théorème 1.11 et du corollaire 1.1 ; ces opérations sont appelées "opérations syntaxiques binaires". Or les opérations chevron et crochet présentées dans [BoNi1, BCN] conduisent à des résultats similaires bien que ces opérations soient unaires ou ne vérifient aucune propriété semblable au lemme syntaxique (exemple du crochet qui est idempotent).

Dans cette partie nous introduisons un nouveau concept qui nous permet d'unifier ces aspects apparemment contradictoires. Une opération n-aire est dite indépendante d'un opérateur m-aire T' si pour tous langages strictement disjoints $L_1,\ldots,L_n$ et tous cônes rationnels fermés par union $\mathcal{L}_1,\ldots,\mathcal{L}_m$, la condition $t(L_1,\ldots,L_n) \in T'(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ implique qu'il existe $1 \leq i \leq n$ et $1 \leq j \leq m$ tels que $L_i \in \mathcal{L}_j$. Ainsi les opérations binaires syntaxiques de S. Greibach et le chevron sont indépendants de l'opérateur auquel ils sont associées et indépendantes de l'opérateur Rat de fermeture rationnelle. De même le crochet est indépendant de Rat et le produit de deux langages est indépendant du crochet.

Ceci nous permet d'établir alors un théorème d'où découle la non principalité d'un certain nombre de cônes rationnels et de retrouver les résultats de [BCN] et [Gr3].

Le cas particulier des opérations indépendantes de Rat sera plus particulièrement étudié dans la partie E et les opérations indépendantes de l'opérateur associé dans la partie F sous le nom d'opérations de degré 0.

Définition 2.5

Nous dirons que des langages $L_1, L_2, \ldots, L_n$ sont strictement disjoints si $L_i \neq \emptyset$, $1 \leq i \leq n$ et pour tout $i \neq j$, $1 \leq i \leq n$, $1 \leq j \leq n$, $X_{L_i} \cap X_{L_j} = \emptyset$.

Définition 2.6

Soit t une opération n-aire ($n \geq 1$) et T' un opérateur m-aire ($m \geq 1$). Nous dirons que t est indépendante de T' si pour tous cônes rationnels fermés par union $\mathcal{L}_1, \ldots, \mathcal{L}_m$ et tous langages strictement disjoints $L_1, \ldots, L_n$, la condition

$$t(L_1, \ldots, L_n) \in T'(\mathcal{L}_1, \ldots, \mathcal{L}_m)$$

implique qu'il existe i, $1 \leq i \leq n$ et j, $1 \leq j \leq m$ tels que $L_i \in \mathcal{L}_j$.

Si l'opérateur T est associé à l'opération t nous dirons que T est indépendant de T'.

Exemples

1) L'opération de bonne substitution et l'opération t_s présentée dans la section précédente sont indépendantes des opérateurs substitution de $\mathcal{L}_1$ dans $\mathcal{L}_2$ et de l'opérateur Rat.

2) L'opération produit de deux langages est indépendante de l'opérateur crochet et les opérations crochet et chevron sont indépendantes des opérateurs Rat et Rat_0.

3) L'opération $\cup$ d'union de deux langages n'est pas indépendante de l'opérateur $\vee$.

4) Soit $\mathcal{L}_1 \wedge \mathcal{L}_2 = \{L_1 \cap L_2 | L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\}$. L'opération $\cap$ n'est pas indépendante de l'opération $\wedge$.

5) Les opérations binaires syntaxiques de S. Greibach [Gr3] sont des opérations indépendantes de Rat et indépendantes de l'opérateur auquel elles sont associées.

Afin d'établir le théorème de la hiérarchie, donnons tout d'abord deux lemmes :

Lemme 2.5

Soit t_1 une opération n-aire associée à un opérateur (fidèle) T_1 et indépendante d'un opérateur conique (fidèle) m-aire T_2. Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels (fidèles) fermés par union tels que

$$T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n) \subsetneq T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))$$

alors

$$T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)) \subsetneq T_1(T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))).$$

Démonstration

Soit $L \in T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))\setminus T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$.
$T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ est un cône rationnel (fidèle), donc il existe n recopies strictement disjointes de L notées $L_1,\ldots,L_n$ telles que
$L_i \in T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))\setminus T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, $1 \le i \le n$.

Alors $t_1(L_1,\ldots,L_n)$ appartient à $T_1(T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)))$ et si $t_1(L_1,\ldots,L_n)$ appartenait à $T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))$, comme t_1 est indépendant de T_2, il en résulterait qu'il existerait i, $1 \le i \le n$, tel que $L_i \in T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, une contradiction.

■

Lemme 2.6

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1 et indépendant de l'opérateur conique (fidèle) m-aire T_2. Si $\mathcal{L}_1,\ldots,\mathcal{L}_m$ sont des cône rationnels (fidèles) fermés par union tels que

$$\mathcal{L}_1 \cup \mathcal{L}_2 \cup \ldots \cup \mathcal{L}_m \subsetneq T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$$

alors $$T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \subsetneq T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)).$$

Démonstration

Soit $L \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \setminus \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_m$ et $L_1,\ldots,L_n$ n recopies strictement disjointes de L. Pour tout i, $1 \leq i \leq n$, $L_i \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \setminus \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_m$ et $t_1(L_1,\ldots,L_n) \in T_1(T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m))$. Supposons que $t_1(L_1,\ldots,L_n) \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$, alors comme t_1 est indépendant de T_2, il existe i, $1 \leq i \leq n$ et j, $1 \leq j \leq n$, tels que $L_i \in \mathcal{L}_j \subseteq \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_m$, une contradiction.

■

On peut alors en déduire le théorème de la hiérarchie :

Théorème 2.2

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1 et t_2 une opération m-aire associée à l'opérateur (fidèle) T_2 tels que t_1 est indépendant de T_2 et t_2 est indépendant de T_1. Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels (fidèles) fermés par union tels que $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n \subsetneq T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, alors le plus petit cône rationnel (fidèle) fermé par union, T_1 et T_2 et contenant $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n$ est non principal.

Démonstration

Soit $\mathcal{E}_p$, $p \geq 1$, la famille des cônes rationnels (fidèles) fermés par union définie par :

$$\mathcal{E}_1 = \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n$$

et si $p \geq 1$ $$\mathcal{E}_{p+1} = T_2(T_1(\mathcal{E}_p)).$$

Montrons que $\mathcal{E}_p \subsetneq \mathcal{E}_{p+1}$. Pour $p = 1$ la condition $\mathcal{E}_1 = \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n \subsetneq T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, implique d'après le lemme 2.6 que $T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n) \not\subseteq T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n))$. Donc du lemme 2.5, il résulte que $\mathcal{E}_2 = T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)) \subsetneq T_1(T_2(T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)))$. Ainsi $\mathcal{E}_1 \subsetneq \mathcal{E}_2$ et $\mathcal{E}_2 \subsetneq T_1(\mathcal{E}_2)$. La preuve s'achève alors par une récurrence simple.

Soit alors $\mathcal{E} = \bigcup_{p \geq 1} \mathcal{E}_p$. $\mathcal{E}$ est un cône rationnel (fidèle) fermé par union, T_1 et T_2. De plus tout cône rationnel (fidèle) contenant $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n$ fermé par union, T_1 et T_2 qui contient $\mathcal{E}_p$ contient $\mathcal{E}_{p+1}$. Donc $\mathcal{E}$ est le plus petit cône rationnel (fidèle) fermé par union, T_1 et T_2 et contenant $\mathcal{L}$ et $\mathcal{E}$ est non principal d'après le théorème 1.8.

∎

Exemples

1) Si l'on prend $T_1 = T_2$ = l'opérateur de substitution on retrouve le corollaire 1.3.

2) Soit $\odot$ l'opérateur associé au produit de deux langages, défini par

$$\mathcal{L}_1 \odot \mathcal{L}_2 = \{L_1L_1' \cup L_2L_2' \cup \ldots \cup L_nL_n' \mid n \geq 1,\ L_i \in \mathcal{L}_1,\ L_i' \in \mathcal{L}_2,\ 1 \leq i \leq n\}.$$

On a vu que le produit de deux langages est indépendant de l'opérateur crochet et que le crochet d'un langage est indépendant de $\odot$ (puisque $\mathcal{L} \odot \mathcal{L} \subseteq \underline{\text{Rat}}\ \mathcal{L}$). Le théorème 2.2 nous permet donc de retrouver le théorème 1.12.

Etablissons maintenant une condition nécessaire et suffisante pour que $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ soit fermé pour une opération t_1 indépendante de T_2 :

Théorème 2.3

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1 et indépendante de l'opérateur conique simple (fidèle) m-aire T_2. Soient $\mathcal{L}_1,\ldots,\mathcal{L}_m$ des cônes rationnels (fidèles) fermés par union. $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ est fermé par t_1 si et seulement si $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) = \mathcal{L}_1 \vee \mathcal{L}_2 \vee \ldots \vee \mathcal{L}_m = T_1(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m)$ et dans ce cas $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ est fermé par T_2.

Démonstration

La condition suffisante est évidente. Réciproquement supposons que $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ est fermé par t_1. Comme T_2 est simple, $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m \subseteq T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ et donc $T_1(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m) \subseteq T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$. Supposons qu'il existe $L \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \setminus T_1(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m)$, alors soient $L_1,\ldots,L_n$ des recopies strictement disjointes de L. L_i appartient à $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \setminus T_1(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m)$, $1 \leq i \leq n$ et donc $t_1(L_1,\ldots,L_n)$ appartient à $T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$. Comme t_1 est indépendant de T_2 il existe i, $1 \leq i \leq n$ et j, $1 \leq j \leq m$ tels que $L_i \in \mathcal{L}_j \subseteq \mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m \subseteq T_1(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m)$, une contradiction.

De même s'il existe un langage $L \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m) \setminus \mathcal{L}_1 \vee \ldots \vee \mathcal{L}_m$,

par un procédé analogue on aboutit à une contradiction.

■

Comme l'opération produit de deux langages est indépendante de l'opérateur crochet, on peut énoncer :

Corollaire 2.7

Si $\mathcal{L}$ est un cône rationnel alors $C(\mathcal{L})$, le crochet de $\mathcal{L}$, est fermé par produit si et seulement si $\mathcal{L} = C(\mathcal{L})$ et $\mathcal{L}$ est fermé par produit.

Nous terminerons cette section en donnant une condition nécessaire pour qu'un opérateur T_1 associé à l'opération t_1 soit indépendant d'un opérateur T_2. Nous montrerons que si t_1 est régulière alors cette condition est suffisante.

Lemme 2.7

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1. Soient $L_1,\ldots,L_n$ des langages strictement disjoints et $\mathcal{L}, \mathcal{L}_1,\ldots,\mathcal{L}_p$ des cônes rationnels (fidèles) tels que la condition $t_1(L_1,\ldots,L_n) \in \mathcal{L}$ implique qu'il existe i, $1 \leq i \leq n$ et j, $1 \leq j \leq p$ tels que $L_i \in \mathcal{L}_j$. Alors pour tous cônes rationnel (fidèles) $\mathcal{C}_1,\ldots,\mathcal{C}_n$ tels que $T_1(\mathcal{C}_1,\ldots,\mathcal{C}_n) \subseteq \mathcal{L}$, il existe i, $1 \leq i \leq n$ tel que $\mathcal{C}_i \subseteq \bigcup_{j=1}^{p} \mathcal{L}_j$. Si de plus t_1 est régulière cette condition est suffisante.

Démonstration

Supposons que pour tout i, $1 \leq i \leq n$, on ait $\mathcal{C}_i \not\subseteq \bigcup_{j=1}^{p} \mathcal{L}_j$, soit alors $L_i \in \mathcal{C}_i \setminus \bigcup_{j=1}^{p} \mathcal{L}_j$, $1 \leq i \leq n$. Comme $\mathcal{C}_i$ est un cône rationnel (fidèle)

on peut supposer que les langages L_i sont strictement disjoints. Alors le langage $t_1(L_1,\ldots,L_n) \in \mathcal{L}$ et l'on aboutit à une contradiction.

Supposons maintenant que t_1 est une opération régulière. Soient $L_1,\ldots,L_n$ des langages strictement disjoints et $\mathcal{L}$ un cône rationnel tels que $t_1(L_1,\ldots,L_n) \in \mathcal{L}$. Alors comme t_1 est régulière $t_1(L_1,\ldots,L_n) \in \mathcal{C}(t_1(L_1,\ldots,L_n)) = T_1(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n)) \subseteq \mathcal{L}$.

Notre hypothèse implique donc qu'il existe i, $1 \leq i \leq n$ tel que $\mathcal{C}(L_i) \subseteq \bigcup_{j=1}^{p} \mathcal{L}_j$ soit $L_i \in \bigcup_{j=1}^{n} \mathcal{L}_j$ i.e. il existe j, $1 \leq j \leq p$ tel que $L_i \in \mathcal{L}_j$, ce qui prouve la réciproque quand t_1 est régulière.

∎

Corollaire 2.8

Soit t_1 une opération (régulière) n-aire associée à l'opérateur (fidèle) T_1 et T_2 un opérateur m-aire conique (fidèle). Une condition nécessaire (et suffisante) pour que t_1 soit indépendante de T_2 est que pour tous cônes rationnels (fidèles) $\mathcal{L}_1,\ldots,\mathcal{L}_n$ et tous cônes rationnels (fidèles) fermés par union $\mathcal{C}_1,\ldots,\mathcal{C}_m$ la condition $T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n) \subseteq T_2(\mathcal{C}_1,\ldots,\mathcal{C}_m)$ implique qu'il existe i, $1 \leq i \leq m$, tel que $\mathcal{L}_i \subseteq \bigcup_{j=1}^{m} \mathcal{C}_j$.

Ainsi dans le cas de l'opérateur crochet, si $\odot$ est l'opérateur associé au produit (cf exemple du chapitre 2), on déduit de la proposition 1.2 :

Corollaire 2.9

Soient $\mathcal{L}_1$, $\mathcal{L}_2$ et $\mathcal{L}_3$ des cônes rationnels, $\mathcal{L}_3$ en outre fermé par union. Si $\mathcal{L}_1 \odot \mathcal{L}_2 \subseteq C(\mathcal{L}_3)$ alors $\mathcal{L}_1$ ou $\mathcal{L}_2$ est inclus dans $\mathcal{L}_3$.

Si l'on prend $\mathcal{L}_3$ = Rat, on sait [BCN] que C(Rat) est le cône rationnel des langages linéaires engendré par $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$.

On retrouve ainsi

Corollaire 2.10 [Gr4]

Rat est le seul cône rationnel fermé par produit contenu dans les langages linéaires.

Ultérieurement nous montrerons par le même procédé que Rat est le seul cône rationnel fermé par produit contenu dans le cône rationnel engendré par $L_1 = \{wcw | w \in \{a,b\}^*\}$.

E) OPERATIONS INDEPENDANTES DE Rat

Nous nous intéressons ici au cas particulier des opérations indépendantes de Rat. Les théorèmes 2.2 et 2.3 de la section précédente nous donnent d'intéressants corollaires qui généralisent à ces opérations des résultats établis par S. Greibach [Gr3] dans le cas des opérations syntaxiques binaires. Ceci nous permet entre autre de présenter une nouvelle version d'une conjecture de S. Greibach sur les FAL "uniformly star closed" [Gr3], i.e. telles que si $\mathcal{L} = \mathcal{F}(\mathcal{S})$ alors $\mathcal{L} = \mathcal{E}(\mathcal{S})$. Nous établissons alors une condition nécessaire et suffisante pour que certaines FAL (fidèles) soient principales (théorème 2.4) ce qui nous permet de simplifier les preuves de résultats connus et d'établir la non principalité de nouvelles FAL. En particulier ce théorème nous permet de montrer que, contrairement à ce qui est énoncé dans [BCN], la FAL engendrée par les langages algébriques bornés est strictement incluse dans la plus petite FAL fermée par chevron.

Nous montrons enfin que Rat et Rat_0 sont les seules FAL fidèles fermées par une opération indépendante de Rat et contenues dans la FAL engendrée par les langages bornés. Ce résultat généralise ceux de Goldstine [Gol] et Greibach [Gr3] respectivement pour la substitution et les opérateurs syntaxiques binaires. Par un procédé analogue nous montrons que la FAL engendré par les langages linéaires ne contient pas de cône rationnel non trivial fermé par produit et par une opération indépendante de Rat (Rat_0).

Des théorèmes 2.2 et 2.3 on tire immédiatement :

Corollaire 2.11

Soit t une opération n-aire associée à l'opérateur (fidèle) T et indépendante de Rat (Rat_0) si le produit de deux langages (resp. l'étoile propre d'un langage) est indépendant de T, alors pour tous cônes rationnels (fidèles) $\mathcal{L}_1, \ldots, \mathcal{L}_n$ tels que $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n \neq T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$, le plus petit cône rationnel (fidèle) fermé par union, produit et T (resp. la plus petite FAL (fidèle) fermée par T) et contenant $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n$ est non principal.

Corollaire 2.12

Soit t une opération n-aire associée à l'opérateur (fidèle) T et indépendante de Rat (resp. Rat_0). Si $\mathcal{L}$ un cône rationnel (fidèle) tel que $\mathcal{F}(\mathcal{L})$ ($\mathcal{F}^f(\mathcal{L})$) est fermé par T, alors :

1) $\mathcal{L} = \mathcal{F}(\mathcal{L}) = \mathcal{F}_T(\mathcal{L})$ (resp. $\mathcal{L} = \mathcal{F}^f(\mathcal{L}) = \mathcal{F}^f_T(\mathcal{L})$)

2) Pour toute famille de langages $\mathcal{S}$, $\mathcal{L} = \mathcal{F}(\mathcal{S})$ si et seulement si $\mathcal{L} = \mathcal{C}^U(\mathcal{S})$ (resp. $\mathcal{L} = \mathcal{F}^f(\mathcal{S})$ si et seulement si $\mathcal{L} = \mathcal{C}^{fU}(\mathcal{S})$).

Démonstration

1) Si $\mathcal{F}(\mathcal{L}) = \text{Rat}\ \mathcal{C}(\mathcal{L})$ est fermé par T, alors d'après le théorème 2.3 $\mathcal{F}(\mathcal{L}) = \text{Rat}\ \mathcal{L} = \mathcal{L}$ et $\mathcal{L}$ est fermé par T.

2) Si $\mathcal{L} = \mathcal{F}(\mathcal{S}) = \text{Rat}\ \mathcal{C}^U(\mathcal{S})$, alors le point 1) implique que $\mathcal{L} = \mathcal{C}^U(\mathcal{S})$.

La preuve est identique pour les opérateurs fidèles.

■

Dans [Gr3] les FAL vérifiant la propriété 2) du corollaire précédent sont dites "uniformly star closed" et S. Greibach pose le problème de savoir si toute FAL possédant cette propriété est nécessairement close pour une opération binaire syntactique au sens de [Gr3]. Le corollaire 2.12 nous permet de simplifier cette conjecture sous la forme

Conjecture 1

Soit $\mathcal{L}$ une FAL, si pour toute famille de langages $\mathcal{S}$ telle que $\mathcal{L} = \mathcal{F}(\mathcal{S})$ on a $\mathcal{L} = \mathcal{C}^U(\mathcal{S})$, alors $\mathcal{L}$ est fermé pour une opération indépendante de Rat.

Nous utiliserons le corollaire 2.12 enfin de cette section pour étudier la FAL engendrée par les langages bornés et celle engendrée par les langages linéaires.

Auparavant établissons un théorème sur la principalité :

Théorème 2.4 [Rod2]

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par union et par une opération n-aire t indépendante de Rat (Rat_0). Alors $\mathcal{L}$ est principal si et seulement si la FAL (fidèle) $\mathcal{F}(\mathcal{L})$ ($\mathcal{F}^f\mathcal{L}$) est principale.

Démonstration

Si $\mathcal{L}$ est un cône rationnel principal, on sait que la FAL $\mathcal{F}(\mathcal{L})$ est principale (théorème 1.7).

Réciproquement si $\mathcal{F}(\mathcal{L}) = \mathcal{F}(L_0)$ pour quelque $L_0 \in \mathcal{F}(\mathcal{L})$, alors puisque $\mathcal{L}$ est fermé par union, il existe un langage $M \in \mathcal{L}$ tel que $L_0 \in \mathcal{F}(M)$. Soit alors L un langage quelconque de $\mathcal{L}$, $L_1,\ldots,L_n$ des copies strictement disjointes de L. Comme $\mathcal{L}$ est fermé par t, on a :

$$t(L_1,\ldots,L_n) \in \mathcal{L} \subseteq \mathcal{F}(\mathcal{L}) = \text{Rat}\ \mathcal{E}(M)$$

et comme t est indépendante de Rat, il existe i, $1 \leq i \leq n$, tel que $L_i \in \mathcal{E}(M)$, i.e. $L \in \mathcal{E}(M)$ et comme $M \in \mathcal{L}$ on a $\mathcal{L} = \mathcal{E}(M)$. La preuve est identique dans le cas des cônes fidèles et de Rat_0.

■

Applications

1) Cônes rationnels commutativement clos :

Soit $X = \{a_1,\ldots,a_p\}$ un alphabet. Pour tout $w \in X^*$ soit $|w|_{a_i}$ le

nombre d'occurrences de la lettre a_i dans w, l'application ψ de X^* dans $\mathbb{N}^p$ est une fonction de Parikh si pour tout $w \in X^*$: $\psi(w) = (|w|_{a_1},\ldots,|w|_{a_p})$.

Etant donné une fonction de Parikh ψ quelconque, la fermeture commutative d'un langage L est l'ensemble $C(L) = \psi^{-1}o\psi(L)$ autrement dit l'ensemble des mots que l'on peut obtenir par permutation des lettres à partir des mots de L. On dit alors qu'une famille de langages $\mathcal{L}$ est commutativement close si $C(L) \in \mathcal{L}$ pour tout $L \in \mathcal{L}$. Dans [Lat3], M. Latteux établit tout d'abord que $\mathcal{L}_\psi$ le plus petit cône rationnel commutativement clos est non principal et fermé par insertion, i.e. l'opération i définie par :

$$i(L_1,L_2) = \{xfy | xy \in L_1, f \in L_2\}.$$

M. Latteux établit alors, par une argumentation assez complexe, que Rat $\mathcal{L}_\psi$ est une FAL non principale. Or, [Rod3], ce résultat découle directement du théorème 2.4 puisque l'opération insertion est indépendante de Rat [Gr3].

2) Langages ultralinéaires et langages fortement quasi-rationnels :

Boasson, Crestin et Nivat montrent dans [BCN] que Ult, la famille des langages ultralinéaires définie dans [GS2] et Qrf la famille des langages fortement quasi-rationnels définis dans [Crest] sont respectivement le plus petit cône rationnel fermé par union, produit, crochet et la plus petite FAL fermée par crochet. Ils démontrent en outre que Ult et Qrf sont non principaux. Le théorème 2.4 permet donc de conclure que Rat Ult est une FAL non principale. Comme de plus Ult n'est pas une FAL, le corollaire 2.12 implique que Rat Ult n'est pas fermé par crochet. D'où :

Proposition 2.1 [Rod2]

1) Rat Ult est une FAL non principale et non fermée par crochet.

2) Ult $\subsetneq$ Rat Ult $\subsetneq$ Qrf.

Dans le même article [BCN], Boasson, Crestin et Nivat démontrent que

$\mathcal{C}_{Ch}^{U\bullet}$(Rat) le plus petit cône rationnel fermé par produit, union et chevron et $\mathcal{F}_{Ch}$(Rat) la plus petite FAL fermée par chevron sont non principaux. Du théorème 2.4 il résulte donc que Rat $\mathcal{C}_{Ch}^{U\bullet}$(Rat) est une FAL non principale. D'autre part nous pouvons établir le lemme :

Lemme 2.8

Soit T un opérateur n-aire (fidèle) conique tel que le produit de deux langages est indépendant de T. Si L est un langage, c un symbole n'appartenant pas à X_L et $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnels (fidèles) fermés par union tels que $(Lc)^+ \in T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, alors il existe i, $1 \leq i \leq n$, tel que $(Lc)^+ \in \mathcal{L}_i$, i.e. l'étoile propre d'un langage est indépendante de T.

Démonstration

Si $(Lc)^+ \in T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, soit L' une recopie de L et d un symbole n'appartenant pas à $X_L \cup X_{L'} \cup \{c\}$. Comme $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ est un cône rationnel, le langage $(L'd)^+(Lc)^+$ appartient à $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ et comme le produit de deux langages est indépendant de T, il en résulte que $(L'd)^+$ ou $(Lc)^+$ appartient à un certain $\mathcal{L}_i$ pour $1 \leq i \leq n$, ce qui prouve le lemme.

■

On en déduit alors immédiatement :

Théorème 2.5

Soit T une opération n-aire conique (fidèle) tel que le produit de deux langages est contraint par T et soient $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnels (fidèles) fermés par union. Une condition nécessaire pour que $\mathcal{C}_T^{U\bullet}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$, le plus petit cône rationnel fermé par union, produit et T et contenant $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n$ (resp. $\mathcal{C}_T^{fU\bullet}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$) soit une FAL est que $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ soit une FAL. Si de plus T est associé

à une opération t indépendante de Rat (Rat_0) alors cette condition est suffisante.

On peut alors énoncer :

Proposition 2.2 [Rod2]

1) Rat $\mathcal{E}^{U.}_{Ch}$ (Rat) est une FAL non principale.

2) $\mathcal{E}^{U.}_{Ch}$ (Rat) $\subsetneq$ Rat $\mathcal{E}^{U.}_{Ch}$ (Rat) $\subsetneq$ $\mathcal{F}_{Ch}$(Rat).

Démonstration

Comme Ch(Rat) = $\mathcal{E}(\{a^n b^n | n \in \mathbb{N}\})$ n'est pas une FAL il en résulte que $\mathcal{E}^{U.}_{Ch}$(Rat) $\subsetneq$ Rat $\mathcal{E}^{U.}_{Ch}$ (Rat) et donc, d'après le corollaire 2.12, Rat $\mathcal{E}^{U.}_{Ch}$(Rat) n'est pas fermé par chevron donc Rat $\mathcal{E}^{U.}_{Ch}$(Rat) $\subsetneq$ $\mathcal{F}_{Ch}$(Rat).

■

Remarque

Dans [Crest2] J.P. Crestin identifie le cône rationnel $\mathcal{E}^{U.}_{Ch}$(Rat) au cône rationnel engendré par les langages algébriques bornés. Rat $\mathcal{E}^{U.}_{Ch}$(Rat) est donc la FAL engendrée par ces langages. Or dans [BCN] Boasson, Crestin et Nivat écrivent que la plus petite FAL fermée par chevron est exactement la FAL engendrée par les langages algébriques bornées étudiée par Goldstine dans [Gold2]. La proposition précédente établit que cette affirmation est fausse.

Nous verrons d'ailleurs, par la suite (proposition 2.3) que Rat est la seule FAL fermée pour une opération indépendante de Rat contenue dans la FAL engendrée par une famille de langages bornés.

En fait de la proposition 2.2 on déduit sans difficulté que le langage $Ch(\{a^nb^n | n \in \mathbb{N}\}^*)$ n'appartient pas à la FAL engendrée par les langages bornés algébriques. Dans le paragraphe suivant nous verrons qu'il n'appartient pas à la FAL engendrée par tous les langages bornés.

3) Langages bornés

Définition 2.7

Un langage L est borné de degré m s'il existe des mots $w_1,\ldots,w_m$ tels que $L \subseteq w_1^* \ldots w_m^*$. Un langage est borné s'il est borné de degré m pour quelque $m \in \mathbb{N}$. Une famille de langages $\mathcal{L}$ est bornée (de degré m) si tous ses éléments sont bornés (de degré m).

Un cône rationnel (fidèle) $\mathcal{L}$ (resp. une FAL (fidèle)) est borné (de degré m) s'il existe une famille de langages (bornée de degré m) $\mathcal{S}$ telle que $\mathcal{L} = \mathcal{E}(\mathcal{S})$ (resp. $\mathcal{E}^f(\mathcal{S})$, $\mathcal{F}(\mathcal{S})$, $\mathcal{F}^f(\mathcal{S})$).

Rappelons deux résultats fondamentaux de [Gr3] :

Lemme 2.9 [Gr3]

Soit $L_i \subseteq X^*$, $1 \leq i \leq n+1$ et $c \notin T$. Soit L un langage inclus dans $a_1^* \ldots a_m^*$. Si $L' = L_1 c\, L_2 c \ldots L_{m+1} c$ appartient à $\mathcal{E}(L)$, alors il existe j, $1 \leq j \leq m+1$ tel que $L_j \in Rat$.

Il en résulte alors directement :

Théorème 2.6 [Gr3]

Soit $\mathcal{S}$ une famille de langages bornés.

1) Si $\mathcal{S}$ est bornée de degré m, $m \geq 1$, alors les seuls cônes rationnels fidèles fermés par produit contenus dans $\mathcal{E}(\mathcal{S})$ sont Rat et Rat_0.

2) Les seules FAL fidèles contenues dans $\mathcal{E}(\mathcal{S})$ sont Rat et Rat_0.

3) $\mathcal{F}(\mathcal{S}) = \mathcal{E}(\mathcal{S})$ si et seulement si $\mathcal{S} \subseteq$ Rat.

Nous pouvons alors énoncer :

Proposition 2.3

Soit $\mathcal{S}$ une famille de langages bornés, les seules FAL fidèles contenues dans $\mathcal{F}(\mathcal{S})$ et fermées pour une opération indépendante de Rat sont Rat et Rat_0.

Démonstration

Si $\mathcal{L}$ est une FAL fidèle contenue dans $\mathcal{F}(\mathcal{S})$ et fermée pour une opération t indépendante de Rat_0, alors pour tout langage $L \in \mathcal{L}$, $t(L_1,\ldots,L_n) \in \mathcal{L}$ pour des copes strictement disjointes de L et puisque $t(L_1,\ldots,L_n) \in$ Rat $\mathcal{E}(\mathcal{S})$ on en déduit que $L \in \mathcal{E}(\mathcal{S})$ i.e. $\mathcal{L} \subseteq \mathcal{E}(\mathcal{S})$. D'après le point 2) du théorème précédent, il en résulte que $\mathcal{L}$ = Rat ou Rat_0.

∎

Corollaire 2.13

Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des FAL fidèles non triviales (i.e. $Rat_0 \neq \mathcal{L}_i$, $1 \leq i \leq n$) et t une opération n-aire indépendante de Rat associée à l'opérateur T, alors $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ n'est pas inclus dans la FAL engendrée par les langages bornés.

Démonstration

Soit $\mathcal{B}$ la famille des langages bornés, si $T(\mathcal{L}_1,\ldots,\mathcal{L}_n) \subseteq \mathcal{F}(\mathcal{B})$ = Rat $\mathcal{E}(\mathcal{B})$ alors d'après le corollaire 2.8, il existe i, $1 \leq i \leq n$ tel que la FAL fidèle non triviale $\mathcal{L}_i$ est incluse

dans $\mathcal{E}(\mathcal{B})$, ce qui contredit le théorème 2.6. ∎

Exemples

1) On sait [Dur] que la famille des langages linéaires, engendrée par $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ n'est pas incluse dans le cône rationnel engendré par les langages bornés. Comme le chevron de S_2, $Ch(S_2)$, est rationnellement équivalent à S_2, il en résulte [Lat3] que S_2 n'appartient pas à la FAL engendrée par les langages bornés. En effet si $S_2 \in \underline{Rat}\ \mathcal{E}(\mathcal{B})$ alors $Ch(S_2) \in \underline{Rat}\ \mathcal{E}(\mathcal{B})$ et donc S_2 appartiendrait à $\mathcal{E}(\mathcal{B})$ ce qui est contradictoire.

2) Nous avons vu précédemment que le langage $L = Ch(\{a^nb^n | n \in \mathbb{N}\}^*)$ n'appartient pas à la FAL engendrée par les langages algébriques bornés. Le corollaire 2.13 nous permet de conclure que L n'appartient pas à $\mathcal{F}(\mathcal{B})$, la FAL engendrée par les langages bornés. En effet, sinon, comme le chevron est une opération régulière $Ch(\mathcal{F}(\{a^nb^n | n \in \mathbb{N}\}))$ appartiendrait à $\mathcal{F}(\mathcal{B})$ ce qui serait contradictoire.

Comme le langage $L = Ch(\{a^nb^n | n \in \mathbb{N}\}^*)$ n'appartient pas à la FAL engendrée par le langage S_2 (cf paragraphe suivant) ceci nous donne un nouvel exemple de langage algébrique n'appartenant pas à la FAL des langages bornés.

2) Langages linéaires :

De manière analogue à ce que nous avons fait pour les langages bornés, nous pouvons déduire d'un lemme de [Gr4] que la FAL engendrée par les langages linéaires ne contient pas de cône rationnel non trivial fermé par produit et par une opération indépendante de Rat.

Lemme 2.9 [Gr4]

Soient L_1 et L_2 deux langages et c un nouveau symbole. Si

$L_1 c L_2 \in \mathcal{E}(S_2) = \text{Lin}$, alors L_1 ou L_2 est rationnel.

Démonstration

Si $L_1 c L_2 \in \mathcal{E}(S_2)$ alors si L_2' est une copie de L_2 strictement disjointe de L_2, le langage $L_1 L_2' \in \mathcal{E}(S_2) = C(\text{Rat})$ et d'après la proposition 1.2 [BCN], L_1 ou $L_2 \in$ Rat.

■

Il en résulte alors :

Théorème 2.7 [Gr4]

Rat et Rat_0 sont les seuls cônes rationnels fidèles fermés par produit contenus dans $\mathcal{E}(S_2) = \text{Lin}$.

On peut alors énoncer :

Proposition 2.4

Les seuls cônes rationnels fidèles fermés par union, produit et par une opération indépendante de Rat (Rat_0) contenus dans $\mathcal{F}(S_2) =$ Rat Lin, sont Rat et Rat_0.

Démonstration

Si $\mathcal{E}$ est un cône rationnel fermé par produit et par une opération t indépendante de Rat alors pour tout langage $L \in \mathcal{E}$, si $L_1,\ldots,L_n$ sont des recopies strictement disjointes de L, le langage $t(L_1,\ldots,L_n) \in \mathcal{E}$. Si $\mathcal{E} \subseteq$ Rat Lin il en résulte que $L \in$ Lin et donc $\mathcal{E} \subseteq$ Lin. D'après le théorème précédent, il en résulte que $\mathcal{E}$ est égal à Rat ou Rat_0.

■

Corollaire 2.14

Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels (fidèles) non triviaux fermés par produit et t une opération n-aire indépendante de Rat (Rat_0) associée à l'opérateur (fidèle) T, alors $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ n'est pas inclus dans la FAL engendrée par les langages linéaires.

Démonstration

Si $T(\mathcal{L}_1,\ldots,\mathcal{L}_n) \subseteq$ Rat Lin, alors d'après le corollaire 2.8, il existe i, $1 \leq i \leq n$ tel que $\mathcal{L}_i \subseteq$ Lin, ce qui contredit le théorème 2.7.

■

De manière analogue on peut montrer pour les opérations indépendantes de Fin :

Proposition 2.5

Rat et Rat_0 sont les seuls cônes rationnels fidèles fermés par produit et par une opération indépendante de Fin (Fin_0) et contenus dans le plus petit cône rationnel fermé par produit contenant les langages linéaires.

Exemple

La FAL engendrée par le langage $L_= = \{a^n b^n | n \in \mathbb{N}\}$ est trivialement égale au cône rationnel engendré par $L_=^+$. Comme l'opération chevron est indépendante de Rat et régulière il résulte du corollaire 2.14 que $Ch(L_=^+)$ n'appartient pas à la FAL engendrée par les langages linéaires, car autrement $\mathcal{C}^{U_\bullet}_{Ch}(L_=^+)$ serait inclus dans la FAL des langages linéaires.

Nous verrons dans la suite de ce travail que l'on peut établir des résultats analogues pour la FAL engendrée par le langage $\{wcw | w \in \{a,b\}^*\}$.

F) OPERATIONS DE DEGRE p

En introduisant l'opération chevron, Boasson, Crestin et Nivat soulignent dans [BCN] et [NiBo1] que cette opération unaire a des propriétés analogues à celles des opérations binaires syntactiques de S. Greibach [Gr3] mais que ces résultats "ne peuvent être simplement déduits des résultats très généraux" de [Gr3]. Nous montrons dans cette partie que la non-principalité des cônes rationnels considérés dans [BCN],[NiBo1], [Rod4], [Rod5] et [Gr3] résulte en fait d'une propriété commune aux opérations considérées : ces opérations sont soit indépendantes de l'opérateur auquel elles sont associée (opérations de d°0), soit elles vérifient une propriété beaucoup plus complexe (opération de d°p > 0). Les lemmes syntaxiques de S. Greibach [Gr3] se généralisent alors de façon simple aux opérations de degré $p \geq 0$, quelle que soit l'arité de ces opérations. Les démonstrations sont le plus souvent identíques à celles de [Gr3] pour les opérations n-aires, $n \geq 2$ mais nécessitent quelques modifications pour les opérations unaires.

Dans [Gr3], une opération binaire t associée à l'opérateur T est dite de degré 0 si pour tous cônes rationnels fermés par union $\mathcal{L}_1$ et $\mathcal{L}_2$ et tous langages strictement disjoints L_1 et L_2, la condition $t(L_1,L_2) \in T(\mathcal{L}_1, \mathcal{L}_2)$ implique qu'il existe $1 \leq i \leq 2$ tel que L_i appartient à $\mathcal{L}_i$. Une opération binaire t associée à l'opérateur T est de degré $p > 0$ s'il existe une opération binaire T' de degré $q < p$ telle que la condition $t(L_1,L_2) \in T(\mathcal{L}_1, \mathcal{L}_2)$ implique qu'il existe i tel que $L_i \in T'(\mathcal{L}_1, \mathcal{L}_2)$.

Nous généralisons de façon très naturelle la notion d'opération de degré 0 pour les opérations n-aires, $n \geq 1$. Par contre dans le cas des opérations de degré $p > 0$ nous n'imposerons pas que l'opérateur T' ait la même arité que l'opération t. Nous verrons en effet que cette démarche est fructueuse pour étudier les opérations crochet ouvert et chevron d'ordre i (chapitre 5 de ce travail). Notons enfin que nous serons obligés, pour établir le théorème 2.9 de cette section et certains lemmes de la section suivante, de restreindre cette définition très générale mais ceci ne nuira en rien aux résultats fondamentaux de la section H sur les FAL décomposables.

Définition 2.8

Soit t une opération n-aire ($n \geq 1$) associée à l'opérateur (fidèle) T. t est de degré 0 si pour tous langages strictement disjoints $L_1, L_2, \ldots, L_n$ et tous cônes rationnels (fidèles) fermés par union $\mathcal{L}_1, \ldots, \mathcal{L}_n$, la condition $t(L_1, \ldots, L_n) \in T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$ implique qu'il existe i, $1 \leq i \leq n$, tel que $L_i \in \mathcal{L}_i$.

Remarque

Toute opération de degré 0 est donc indépendante de l'opérateur T qui lui est associé. Les résultats des sections précédentes sont donc vérifiés par ces opérations.

Il faut noter, par ailleurs, que les preuves des théorèmes de cette section établis dans le cas des opérations de degré 0 restent vraies dans le cas général des opérations indépendantes de leur opérateur associés i.e. tels que la condition $t(L_1, \ldots, L_n) \in T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$ implique qu'il existe i, $1 \leq i \leq n$ et j, $1 \leq j \leq n$ tels que $L_i \in \mathcal{L}_j$.

Exemples

0) L'opération produit de deux langages est de degré 0 [Gr3].

1) L'opération unaire chevron est de degré 0 [Lat4].

2) L'opération unaire crochet n'est pas de degré 0, puisque l'opérateur associé est idempotent, i.e. $C(\mathcal{L}) = C(C(\mathcal{L}))$.

3) L'opération substitution est de degré 0 [Gr2].

4) Rappelons la définition de l'opération insertion introduite dans [Gr3].

Définition 2.8

Soient L_1 et L_2 deux langages, l'insertion de L_2 dans L_1 est l'opération binaire définie par

$$i(L_1,L_2) = \{xfy \mid xy \in L_1, f \in L_2\}.$$

I désignera l'opérateur associé à l'opération insertion.

S. Greibach montre dans [Gr3] que l'insertion n'est pas de degré 0.

Nous allons établir maintenant un lemme qui permet de conclure que des opérations sont de degré 0. Définissons tout d'abord :

Définition 2.9

Soit T_1 un opérateur (fidèle) n-aire, $n \geq 1$, et T_2 un opérateur (fidèle) m-aire avec $m \leq n$. Nous dirons que T_2 est inclus dans T_1, et nous noterons $T_2 \subseteq T_1$ si et seulement si pour tous cônes rationnels (fidèles) fermés par union $\mathscr{L}_1,\ldots,\mathscr{L}_n$ on a

$$T_2(\mathscr{L}_{j_1},\ldots,\mathscr{L}_{j_m}) \subseteq T_1(\mathscr{L}_1,\ldots,\mathscr{L}_n)$$

pour tous $1 \leq j_1 < \ldots < j_m \leq n$.

On a alors :

Lemme 2.10

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1 et t_2 une opération m-aire associée à l'opérateur (fidèle) T_2. Si T_1 est inclus dans T_2 et si pour tous langages strictement disjoints $L_1,\ldots,L_n$ et tous cônes rationnels (fidèles) fermés par union $\mathscr{L}_1,\ldots,\mathscr{L}_m$, la condition $t_1(L_1,\ldots,L_n) \in T_2(\mathscr{L}_1,\ldots,\mathscr{L}_m)$ implique qu'il existe i, $1 \leq i \leq n$, tel que

$L_i \in \mathcal{L}_i$, alors t_1 est de degré 0. Si de plus t_2 est régulière, alors t_2 est aussi de degré 0.

Démonstration

1) Soient $L_1,\ldots,L_n$ des langages strictement disjoints, $\mathcal{L}_1,\ldots,\mathcal{L}_n$ n cônes rationnels (fidèles) fermés par union, tels que $t_1(L_1,\ldots,L_n) \in T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$. Comme $T_1 \subseteq T_2$ il en résulte que $t_1(L_1,\ldots,L_n) \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_n,\ldots,\mathcal{L}_n)$ et la condition de ce lemme implique qu'il existe i, $1 \leq i \leq n$ tel que $L_i \in \mathcal{L}_i$, i.e. t_1 est de degré 0.

2) Soient $L_1,\ldots,L_m$ des langages strictement disjoints, $\mathcal{L}_1,\ldots,\mathcal{L}_m$ des cônes rationnels (fidèles) fermés par union, tels que

$$t_2(L_1,\ldots,L_m) \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m).$$

Si t_2 est une opération régulière alors $\mathcal{C}(t_2(L_1,\ldots,L_m)) = T_2(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_m)) \subseteq T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ et comme $T_1 \subseteq T_2$ il en résulte que $T_1(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_n)) \subseteq T_2(\mathcal{C}(L_1),\ldots,\mathcal{C}(L_m))$ soit donc que $t_1(L_1,\ldots,L_n) \in T_2(\mathcal{L}_1,\ldots,\mathcal{L}_m)$ et l'hypothèse du lemme implique qu'il existe i, $1 \leq i \leq n$, tel que $L_i \in \mathcal{L}_i$, i.e. que t_2 est de degré 0.

■

Il en résulte alors immédiatement :

Corollaire 2.15

Si t_1 et t_2 sont deux opérations unaires associées respectivement aux opérateurs T_1 et T_2, si t_2 est régulière, $T_1 \subseteq T_2$ et t_1 est indépendante de T_2, alors t_1 et t_2 sont de degré 0.

Afin de définir les opérations de degré $p > 0$, introduisons la notion d'opérateur supérieur ou égal à un autre opérateur :

Définition 2.10

Soit t_1 une opération n-aire associée à un opérateur (fidèle) T_1. Nous dirons que T_1 est supérieur ou égal à l'opérateur (fidèle) m-aire T_2 et nous noterons $T_1 \geq T_2$ si :

1) $T_2 \subseteq T_1$.

2) Si $L_1,\ldots,L_n$ sont des langages strictement disjoints, $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnels (fidèles) fermés par union, alors la condition $t_1(L_1,\ldots,L_n) \in T_1(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ implique qu'il existe i, $1 \leq i \leq n$ $1 \leq j_1 < \ldots < j_m \leq n$ tels que

$$L_i \in T_2(\mathcal{L}_{j_1},\ldots,\mathcal{L}_{j_m})$$

Définition 2.11

Une opération t associée à l'opérateur T est de degré $p \geq 1$ si il existe m, $0 \leq m < p$, tel que T est supérieur ou égal à un opérateur de degré m.

Exemple

Soit I l'opérateur insertion associé à l'opération de même nom. S. Greibach montre dans [Gr3] que l'opérateur I est supérieur ou égal à l'opérateur produit $\odot$, elle établit en outre que l'opération i n'est pas de degré 0. Ceci montre donc l'existence d'opérations de degré p qui ne sont pas de degré 0.

Remarque

Si T_1 et T_2 sont deux opérateurs tels que $T_1 \geq T_2$, la définition 2.10 implique donc que l'arité de T_2 est inférieure ou égale à l'arité de T_1. Nous étudierons plus loin une hiérarchie d'opérateurs appelés crochets ouverts

d'ordre n, noté COn, qui vérifient COn ≥ CO(n-1) et où l'arité de COn est exactement n. Notons que cette définition se révèlera trop générale dans le cas de l'étude des "FAL décomposables" (Cf paragraphe G) pour l'établissement de certains lemmes, cependant le théorème principal de la section G restera vrai dans le cadre de la définition 2.10.

Etablissons maintenant un lemme d'où découle le théorème de hiérarchie qui généralise les résultats de [Gr3] :

Lemme 2.11

Soit t une opération n-aire de degré p ≥ 0, associée à l'opérateur (fidèle) T. Si $\mathcal{L}$ est un cône rationnel (fidèle) non formé par T alors T($\mathcal{L}$) n'est pas fermé par T.

Démonstration

Procédons par induction sur p.

1) Si p = 0, alors le résultat découle du lemme 2.5 puisque t est indépendant de T.

2) Supposons que p ≥ 1 et que le résultat est établi pour tout m < p. Soit t une opération de degré p, alors T est supérieur ou égal à un opérateur m-aire T' associé à une opération t' de degré q < p. Deux cas sont possibles :

a) Si T'($\mathcal{L}$) = $\mathcal{L}$, alors soient $L_1,\ldots,L_n$ des copies strictement disjointes de $L \in T(\mathcal{L}) \setminus \mathcal{L}$. Alors pour tout i, $1 \leq i \leq n$, $L_i \in T(\mathcal{L}) \setminus \mathcal{L}$ et $t(L_1,\ldots,L_n) \in T(T(\mathcal{L}))$. Supposons que $t(L_1,\ldots,L_n)$ appartient à T($\mathcal{L}$), alors il en résulte qu'il existe i, $1 \leq i \leq n$, tel que $L_i \in \mathcal{L}$ ce qui est contradictoire. Donc $T(\mathcal{L}) \neq T(T(\mathcal{L}))$.

b) Supposons alors que $\mathcal{L} \neq T'(\mathcal{L})$. D'après l'hypothèse de récurrence T'($\mathcal{L}$) n'est pas fermé par T' et il existe donc n langages strictement disjoints $L_1,\ldots,L_n$ appartenant à $\mathcal{E}^U_{T'}(\mathcal{L}) \setminus \mathcal{L}$. Comme T' ≤ T, on a donc $L_i \in \mathcal{E}^U_T(\mathcal{L})$, $1 \leq i \leq n$ et $t(L_1,\ldots,L_n) \in \mathcal{E}^U_T(\mathcal{L})$. Supposons que

$t(L_1,\ldots,L_n)$ appartient à $T(\mathcal{L})$, alors il en résulte qu'il existe i, $1 \leq i \leq n$, tel que $L_i \in T'(\mathcal{L})$ ce qui est contradictoire. Donc $T(\mathcal{L})$ n'est pas fermé par T.

■

On peut alors énoncer :

Théorème 2.8

Soit t une opération n-aire de degré $p \geq 0$ associée à l'opérateur (fidèle) T et soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par union et non fermé par T. Alors

1) $\forall n \geq 0$, $T^n(\mathcal{L})$ n'est pas fermé par T et $T^n(\mathcal{L}) \not= T^{n+1}(\mathcal{L})$.

2) $\mathcal{E}_T^U(\mathcal{L})$ ($\mathcal{E}_T^{fU}(\mathcal{L})$) n'est pas principal.

Démonstration

On montre par une récurrence simple, à partir du lemme précédent, que $T^n(\mathcal{L}) \not= T^{n+1}(\mathcal{L})$. Le point 2) résulte alors directement du théorème 1.8.

■

Le théorème 2.8 permet de retrouver ainsi les résultats de [NiBo1], [Gr2] ou [Gr3]. En utilisant par ailleurs les propriétés des opérations indépendantes de Rat, il permet d'établir de nouveaux résultats, tels que :

Corollaire 2.16

Soit $\mathcal{L}$ un cône rationnel principal non fermé par produit. Alors $\mathcal{E}^{U\bullet}(\mathcal{L})$ la fermeture par union et produit de $\mathcal{L}$ est un cône rationnel non principal qui n'est fermé pour aucune opération indépendante Rat.

Démonstration

Comme le produit de deux langages est de degré 0, le théorème précédent implique que $\mathcal{E}^{U\bullet}(\mathcal{L})$ est non principal. Si $\mathcal{E}^{U\bullet}(\mathcal{L})$ était fermé par une opération indépendante de Rat alors il résulterait du théorème 2.4 que $\mathcal{F}(\mathcal{L})$ serait non principal, ce qui est contradictoire.

∎

Ainsi le cône rationnel fermé par produit engendré par les langages linéaires n'est fermé pour aucune opération indépendante par Rat (Cf plus haut).

Afin de répondre à un problème ouvert de S. Ginsburg [Gi2,p 242] introduisons une nouvelle définition :

Définition 2.12

Une opération n-aire t associée à l'opérateur T est strictement de degré p s'il existe une suite d'opérateurs n-aires $T_0, T_1, \ldots, T_{p-1}$ telle que $T \geq T_{p-1} \geq \ldots \geq T_0$ et T_0 est associé à une opération de degré 0.

Nous avons alors :

Théorème 2.9

Soit t une opération n-aire, $n \geq 2$, strictement de degré $p \geq 0$ associée à l'opérateur (fidèle) T. Si $\mathcal{L}_1, \ldots, \mathcal{L}_n$ sont des cônes rationnels (fidèles) fermés par union incomparables, alors $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n$ et $T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$ ne sont pas fermés par t.

Démonstration

T est simple donc si $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n$ est fermé par T, alors $T(\mathcal{L}_1, \ldots, \mathcal{L}_n) \subseteq \mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n \subseteq T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$. Il suffit donc de

démontrer que $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ n'est pas fermé par T. Procédons par récurrence sur p.

1) Si p = 0 alors soit $L_i \in \mathcal{L}_i \setminus (\bigcup_{j \neq i} \mathcal{L}_j)$, $1 \leq i \leq n$. Nous pouvons supposer que les langages L_i sont strictement disjoints. Alors :

a) Si n est pair, alors

$t(L_n,L_{n-1},\ldots,L_2,L_1) \in T(\mathcal{L}_n,\mathcal{L}_{n-1},\ldots,\mathcal{L}_2,\mathcal{L}_1)$. Comme T est simple $t(L_n,\ldots,L_1) \in T(T(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ et $t(L_n,\ldots,L_1) \notin T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ puisque cela impliquerait qu'il existerait i, $1 \leq i \leq n$, tel que $L_i \in \mathcal{L}_{n-i+1}$, ce qui est contradictoire.

b) Si $n = 2p+1$, $p > 0$, alors :

$t(L_{2p+1},L_{2p},\ldots,L_{p+3},L_p,L_{p+2},L_{p+1},L_{p-1},\ldots,L_1)$ appartient à $T(T(\mathcal{L}_1,\ldots,\mathcal{L}_n))$ et de même qu'au point a) ce langage ne peut appartenir à $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$.

2) Supposons $p > 0$ et le théorème établit pour tout $m < p$. Si t est strictement de degré p, alors il existe un opérateur n-aire T' de degré $q < p$ tel que $T \geq T'$.

Par hypothèse de récurrence $T'(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ n'est pas fermé pour T'. Soient alors $L_1,\ldots,L_n$ des langages strictement disjoints appartenant à $\mathcal{C}^U_{T'}(\mathcal{L}_1,\ldots,\mathcal{L}_n) \setminus T'(\mathcal{L}_1,\ldots,\mathcal{L}_n)$. Comme $T' \geq T$, les langages L_i appartiennent à $\mathcal{C}^U_T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$. Soit alors $L_0 = t(L_1,\ldots,L_n)$. $L_0 \in \mathcal{C}^U_T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$. Si L_0 appartenait à $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ alors il existerait i, $1 \leq i \leq n$, tel que $L_i \in T'(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, ce qui est contradictoire. Donc $T(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ n'est pas fermé par t.

■

Nous terminerons ce paragraphe en donnant une condition nécessaire et suffisante pour que $T(\mathcal{L})$ et $T(\mathcal{L}')$ soient incomparables, si T est un opérateur associé à une opération de degré 0.

Lemme 2.12

Soit t une opération n-aire de degré 0 associée à l'opérateur (fidèle) T. Soient $\mathcal{L}$ et $\mathcal{L}'$ deux cônes rationnels (fidèles) fermés par union. Alors $T(\mathcal{L})$ est incomparable à $T(\mathcal{L}')$ si et seulement si $\mathcal{L}$ est incomparable à $\mathcal{L}'$.

Démonstration

Si $T(\mathcal{L})$ et $T(\mathcal{L}')$ sont incomparables, alors $\mathcal{L}$ et $\mathcal{L}'$ sont incomparables. En effet, si $\mathcal{L} \subseteq \mathcal{L}'$ alors on aurait $T(\mathcal{L}) \subseteq T(\mathcal{L}')$, une contradiction.

Réciproquement si $\mathcal{L}$ est incomparable à $\mathcal{L}'$, alors soit $L \in \mathcal{L} \setminus \mathcal{L}'$ et $L_1,\ldots,L_n$ des copies strictement disjointes de L. Le langage $t(L_1,\ldots,L_n)$ appartient à $T(\mathcal{L})$ et si $t(L_1,\ldots,L_n)$ appartenait à $T(\mathcal{L}')$ alors comme t est de degré 0, il en résulterait qu'un des L_i appartiendrait à $\mathcal{L}'$, soit $L \in \mathcal{L}'$. D'où le résultat.

■

Exemple

On sait [Berst] que les langages $L_< = \{a^n b^m | 0 \leq n \leq m\}$ et $L_> = \{a^n b^m | n \geq m \geq 0\}$ sont rationnellement incomparables, i.e. $\mathcal{C}(L_<)$ et $\mathcal{C}(L_>)$ sont incomparables. Le lemme 2.12 permet de conclure que les langages $Ch(L_<) = \{c^p a^n b^m d^p | p \in \mathbb{N}, 0 \leq n \leq m\}$ et $Ch(L_>) = \{c^p a^n b^m d^p | p \in \mathbb{N}, n \geq m \geq 0\}$ sont rationnellement incomparables. Cependant il faut noter que puisque $L_= = \{a^n b^n | n \in \mathbb{N}\}$ domine rationnellement $L_>$ et $L_<$ et que $L_=$ appartient à Ch(Rat) on a :

$$\mathcal{C}^U_{Ch}(L_<) = \mathcal{C}^U_{Ch}(L_>) = \mathcal{C}^U_{Ch}(L_=) = \mathcal{C}^U_{Ch}(Rat).$$

G) OPERATIONS SYNTAXIQUES

Une opération syntaxique (ou syntactique) est une opération de degré $p \geq 0$ indépendante de Rat. En rapprochant les résultats des deux sections précédentes on établit donc pour toute opération syntactique n-aire, $n \geq 1$, des propriétés analogues à celles de [Gr3].

Nous définissons alors les opérateurs syntactiques comme des opérateurs associés à une opération syntactique. Nous montrons alors que l'opérateur syntactique insertion peut être associé à une opération de degré un qui n'est pas indépendante de Rat. Ceci répond à une question ouverte de S. Greibach dans [Gr3] et nous conduit à poser une nouvelle conjecture : toute opération associée à un opérateur syntactique est de degré $p \geq 0$.

Définition 2.13

Soit t une opération n-aire, $n \geq 1$, associée à l'opérateur T. On dira que t est une opération syntaxique si t est de degré $p \geq 0$ et t est indépendante de Rat.

Un opérateur (fidèle) T est syntaxique s'il est associé à une opération syntaxique.

Ainsi les opérations chevron, substitution et insertion sont syntaxiques. L'opération produit de deux langages n'est pas syntaxique puisqu'elle n'est pas indépendante de Rat et le crochet n'est pas syntaxique puisqu'il est idempotent.

Le lemme suivant nous permet de conclure, qu'une opération est syntaxique sans vérifier la totalité de la définition 2.13 :

Lemme 2.13

Soit t_1 une opération n-aire associée à l'opérateur (fidèle) T_1 supérieur à un opérateur syntaxique (fidèle) T_2 associé à l'opération m-aire t_2.

Si pour tout langages strictement disjoints $L_1,\ldots,L_m,\ldots,L_n$ on a $t_2(L_1,\ldots,L_m) \in \mathcal{E}(t_1(L_1,\ldots,L_n))$ (resp. $t_2(L_1,\ldots,L_m) \in \mathcal{E}^f(t_1(L_1,\ldots,L_n))$ alors t_1 est syntaxique.

Démonstration

Il suffit de démontrer que t_1 est indépendante de Rat. Soit $\mathcal{L}$ un cône rationnel fermé par union, $L_1,\ldots,L_m,\ldots,L_n$ des langages strictement disjoints tels que $t_1(L_1,\ldots,L_n) \in \underline{\text{Rat}}\ \mathcal{L}$. Comme $t_2(L_1,\ldots,L_m) \in \mathcal{E}(t_1(L_1,\ldots,L_n))$, on a $t_2(L_1,\ldots,L_m) \in \underline{\text{Rat}}\ \mathcal{L}$ et donc il existe i, $1 \leq i \leq m$, tel que $L_i \in \mathcal{L}$ puisque t_2 est syntaxique. La preuve est identique si T_2 est fidèle.

■

Des résultats établis aux paragraphes E et F il résulte alors immédiatement :

Corollaire 2.17

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé pour un opérateur syntaxique (fidèle) n-aire, $n \geq 1$, T. Si $\mathcal{S}$ est une famille de langages telle que $\mathcal{L} \subseteq \mathcal{F}(\mathcal{S})$ ($\mathcal{F}^f(\mathcal{S})$), alors $\mathcal{L} \subseteq \mathcal{E}(\mathcal{S})$ ($\mathcal{E}^f(\mathcal{S})$).

Corollaire 2.18

Soit T un opérateur syntaxique n-aire, $n \geq 1$, (fidèle) et $\mathcal{L}$ un cône rationnel (fidèle) fermé par union tel que $\mathcal{F}(\mathcal{L})$ ($\mathcal{F}^f(\mathcal{L})$) est fermé pour T. Alors

1) $\mathcal{L} = \mathcal{F}(\mathcal{L}) = \mathcal{F}_T(\mathcal{L})$ ($= \mathcal{F}^f(\mathcal{L}) = \mathcal{F}^f_T(\mathcal{L})$).

2) Pour toute famille de langages $\mathcal{S}$, $\mathcal{L} = \mathcal{F}(\mathcal{S})$ si et seulement si $\mathcal{L} = \mathcal{E}^U(\mathcal{S})$ ($\mathcal{L} = \mathcal{F}^f(\mathcal{S})$ si et seulement si $\mathcal{L} = \mathcal{E}^{fU}(\mathcal{S})$).

Théorème 2.10

Soit T un opérateur syntaxique n-aire ($n \geq 1$) et $\mathcal{L}$ un cône rationnel non fermé par T. Alors :

1) $\forall p \geq 0$, $T^p(\mathcal{L})$ n'est pas fermé par T et $T^p(\mathcal{L}) \subsetneq T^{p+1}(\mathcal{L})$.

2) $\mathcal{E}_T^U(\mathcal{L})$ est non principal.

3) $\forall p \geq 0$, $\varphi_T^p(\mathcal{L})$ n'est pas fermé par T et

$\varphi_T^p(\mathcal{L}) \not\subseteq \varphi_T^{p+1}(\mathcal{L})$.

4) $\mathcal{E}_T^{U\bullet}(\mathcal{L})$ est non principal.

5) $\forall p \geq 0$, $\psi_T^p(\mathcal{L})$ n'est pas fermé par T et $\psi_T^p(\mathcal{L}) \subsetneq \psi_T^{p+1}(\mathcal{L})$.

6) $\mathcal{F}_T(\mathcal{L})$ est non principal.

Corollaire 2.19

Si $\mathcal{L}_1$ est un cône rationnel principal et $\mathcal{L}_2$ un cône rationnel tel que $\mathcal{L}_2 \not\subseteq \mathcal{L}_1$ alors $\mathcal{L}_1 \neq \mathcal{E}_T^U(\mathcal{L}_2)$ pour tout opérateur syntaxique T.

Corollaire 2.20

Si $\mathcal{L}$ est un cône rationnel principal, T un opérateur syntaxique et $\mathcal{S}$ une famille de langages telle que $\mathcal{L} = \mathcal{F}_T(\mathcal{S})$, alors $\mathcal{L} = \mathcal{E}^U(\mathcal{S})$.

Corollaire 2.21

Soit T un opérateur syntaxique n-aire, $n > 1$, $\mathcal{L}_1, \ldots, \mathcal{L}_n$ des cônes rationnels incomparables. Alors

1) $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n$, $T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$, Fin $(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$, Fin $T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$, Rat $(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$ et Rat $T(\mathcal{L}_1, \ldots, \mathcal{L}_n)$ ne sont

pas fermés par T.

2) $\mathcal{E}^{U}_{T}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$ est non principal.

3) $\mathcal{E}^{U.}_{T}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$ est non principal.

4) $\mathcal{F}_{T}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$ est non principal.

Exemple

J.P. Crestin étudie dans [Crest2] la famille B_k des cônes rationnels des langages algébriques bornés d'ordre k, définie par

$$B_1 = \mathcal{E}^{U}_{Ch}(Rat),$$

et pour $k > 1$

$$B_k = \mathcal{E}^{U}_{Ch}(\underline{Fin}\ B_{k-1}).$$

Comme Rat n'est pas fermé par chevron, il résulte du théorème 2.10 que B_1 est non principal et comme le chevron est indépendant de Rat (donc de Fin) et que B_1 n'est pas clos par produit, il en résulte que Fin B_1 n'est pas fermé par chevron. Par une récurrence simple on en déduit donc que B_k est non principal et strictement inclus dans B_{k+1}, pour tout $k \geq 1$.

B, le cône rationnel fermé par union et produit engendré par les langages algébriques bornés qui est égal à $\bigcup_{k \geq 1} B_k$ est donc non principal [BCN].

Remarque

Dans [Gr3, p 38], S. Greibach remarque qu'un opérateur syntaxique peut être associé à deux opérations distinctes et pose le problème de savoir si une des opérations peut être syntaxique et l'autre pas. Nous donnons maintenant une réponse affirmative à cette question :

Définition 2.14

Soit $Y \subseteq X$ deux ensembles finis. Un langage $L \subseteq X^*$ est un Y-langage si $L \subseteq (X \setminus Y)^* Y (X \setminus Y)^*$.

Définition 2.15

Soit i_1 l'opération binaire définie par :

1) Si $L_1 \subseteq X^*$ est un Y-langage pour quelque $Y \subseteq X^*$, alors $i_1(L_1,L_2) = \{xfy \mid f \in L_2, \exists c \in Y : xcy \in L_1\}$.

2) Sinon $i_1(L_1,L_2) = L_1 L_2$.

On a alors :

Lemme 2.14

L'opération i_1 est associée à l'opérateur insertion.

Démonstration

De la définition de l'opérateur insertion donnée dans [Gr3] il résulte directement que pour tous cônes rationnels fermés par union $\mathcal{L}_1$ et $\mathcal{L}_2$ on a :

$$\{i_1(L_1,L_2) \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\} \subseteq I(\mathcal{L}_1, \mathcal{L}_2).$$

Réciproquement si $L_1 \in \mathcal{L}_1$ et $L_2 \in \mathcal{L}_2$, alors soit c un nouveau symbole n'appartenant pas à X_{L_1}. Le langage $L_1' = \{xcy \mid xy \in L_1\}$ est rationnellement équivalent à L_1 et L_1' est un $\{c\}$-langage tel que $i(L_1,L_2) = i_1(L_1',L_2)$, d'où le résultat.

■

Or l'opération i_1 n'est pas syntactique. En effet si $\mathcal{L}$ est un cône rationnel fermé par union, non clos par produit, soient L_1 et $L_2 \in \underline{\text{Rat}}\ \mathcal{L} \setminus \mathcal{L}$. Nous pouvons supposer que L_1 n'est pas un Y-langage puisque $\mathcal{L}$ est un cône rationnel. On a alors :

$i_1(L_1,L_2) = L_1L_2 \in \underline{\text{Rat}}\ \mathcal{L}$ mais ni L_1, ni L_2 n'appartient à $\mathcal{L}$.

Donc i_1 n'est pas indépendant de Rat. On peut démontrer cependant que i_1 est de degré 1.

Ceci nous amène donc à poser la

Conjecture 1

Si T est un opérateur syntactique, alors toute opération t associée à T est de degré p, pour quelque $p \in \mathbb{N}$.

H) FAL DECOMPOSABLES

Une FAL $\mathcal{L}$ est décomposable s'il existe deux FAL incomparables $\mathcal{L}_1$ et $\mathcal{L}_2$ telles que $\mathcal{L} = \mathcal{F}(\mathcal{L}_1 \cup \mathcal{L}_2)$. Cette notion, introduite par S. Greibach dans [Gr3] est d'un intérêt primordial pour montrer qu'une FAL contient strictement une FAL maximale. En effet une FAL principale $\mathcal{L} \neq$ Rat est indécomposable si et seulement si la famille des langages non générateurs de $\mathcal{L}$, $N_g(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{F}(L) \neq \mathcal{L}\}$, est l'unique FAL maximale contenue dans $\mathcal{L}$ [Gr3].

S. Greibach établit ainsi que tout FAL principale fermée pour un opérateur syntaxique binaire est indécomposable et recouvre donc une FAL maximale. Les preuves de [Gr3] se généralisent facilement au cas des opérations syntaxiques n-aires pour $n \geq 2$ si ces opérations sont strictement de degré p. Pour traiter le cas des opérations syntactiques de degré p, il nous faut nécessairement considérer des FAL fermées pour des opérations unaires de degré 0 pour lesquelles les preuves de [Gr3] ne sont plus vraies. Pour cela nous établissons (lemme 2.14) que toute FAL fermée par une opération syntaxique unaire est fermée pour une opération syntaxique binaire, ce qui nous permet d'établir le théorème 2.13 qui généralise, dans tous les cas, les résultats de [Gr3].

Les conditions de ce théorème ne sont cependant pas nécessaires. Dans [Gr3], S. Greibach montre que les FAL engendrées par $L_= = \{a^n c b^n \mid n \in \mathbb{N}\}$ et $S_2 = \{wc\tilde{w} \mid w \in \{a,b\}^*\}$ sont indécomposables mais ne sont pas fermées par une opération syntactique. Nous établissons qu'il en est de même pour la FAL engendrée par $L_1 = \{wcw \mid w \in \{a,b\}^*\}$.

Soulignons que si les théorèmes 2.11 et 2.12 sont présentés dans [Gr3] dans le cas des FAL et des FAL fidèles, S. Greibach n'établit le résultat fondamental (théorème 3.3 de [Gr3]) que pour les FAL et pose le problème de savoir si une FAL fidèle principale fermée pour une opération syntactique est indécomposable. Le théorème 3.13 donne une réponse affirmative à cette question dans un cas particulier.

Remarquant alors que le théorème 3.13 reste vrai dans le cas des cônes rationnels fidèles fermés par produit, homomorphisme à effacement linéaire et par une opérationsyntactique, nous en déduisons, par exemple, que $\mathcal{L}_{BNP}$ le plus petit cône rationnel fidèle fermé par intersection contenant $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et $\mathcal{L}_{DUP}$ le plus petit cône rationnel fidèle fermé par intersection contenant $L_1 = \{wcw | w \in \{a,b\}^*\}$ recouvrent chacun strictement un cône rationnel fidèle fermé par chevron.

Rappelons tout d'abord les définitions et notations introduites par S. Greibach dans [Gr3].

Définition 2.16

Soit $\mathcal{L}$ une FAL (fidèle), l'ensemble des langages non générateurs de $\mathcal{L}$ est

$$Ng(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{F}(L) \neq \mathcal{L}\}$$

$$(Ng^f(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{F}^f(L \cup \{\varepsilon\}) \neq \mathcal{F}^f(\mathcal{L} \cup \{\varepsilon\})\}).$$

Définition 2.17

Une FAL (resp. fidèle) $\mathcal{L}$ est (fidèlement) décomposable s'il existe des FAL (fidèles) incomparables $\mathcal{L}_1$ et $\mathcal{L}_2$ (resp. $\mathcal{F}^f(\mathcal{L}_1 \cup \{\varepsilon\})$ est incomparable à $\mathcal{F}^f(\mathcal{L}_2 \cup \{\varepsilon\})$ telles que $\mathcal{L} = \mathcal{F}(\mathcal{L}_1 \cup \mathcal{L}_2)$) ($\mathcal{L} = \mathcal{F}^f(\mathcal{L}_1 \cup \mathcal{L}_2)$).

Définition 2.18

Une FAL (fidèle) $\mathcal{L}$ recouvre une FAL (fidèle) $\mathcal{L}'$ si $\mathcal{L}' \subseteq \mathcal{L}(\mathcal{L}' \cup \{\varepsilon\} \subsetneq \mathcal{L} \cup \{\varepsilon\})$ et il n'existe pas de FAL (fidèle) $\mathcal{L}''$ telle que $\mathcal{L}' \subsetneq \mathcal{L}'' \subsetneq \mathcal{L}(\mathcal{L}' \cup \{\varepsilon\} \subsetneq \mathcal{L}'' \cup \{\varepsilon\} \subsetneq \mathcal{L} \cup \{\varepsilon\})$.

On a alors :

Théorème 2.11 [Gr3]

Soit $\mathcal{L}$ une FAL (fidèle) principale non triviale. Les propositions suivantes sont équivalentes :

1) $\mathcal{N}_g(\mathcal{L})$ est une FAL ($Ng^f(\mathcal{L})$ est une FAL fidèle).

2) $Ng(\mathcal{L})$ ($Ng^f(\mathcal{L})$) est fermé par union.

3) $\mathcal{L}$ n'est pas décomposable.

4) $\mathcal{L}$ n'est pas décomposable en FAL (fidèles) principales.

5) $Ng(\mathcal{L})$ ($Ng^f(\mathcal{L})$) est l'unique FAL (fidèle) recouverte par $\mathcal{L}$.

Corollaire 2.22 [Gr3]

Une FAL (fidèle) principale recouvre une et une seule FAL (fidèle) si et seulement si elle est indécomposable.

Théorème 2.12 [Gr3]

Si $\mathcal{L}$ est une FAL (fidèle) non principale et indécomposable, alors $\mathcal{L}$ ne recouvre aucune FAL (fidèle).

Dans [Gr3] S. Greibach établit les théorèmes 2.11 et 2.12 dans le cas des FAL et mentionne que le résultat reste vrai pour les FAL fidèles. Cependant l'équivalent pour les opérations syntaxiques binaires des théorèmes 2.13 et 2.14 n'est donné que dans le cas des cônes rationnels et des FAL et S. Greibach pose le problème de savoir si ces propriétés restent vraies pour les FAL fidèles. Nous répondons ici affirmativement à cette question et quelle que soit l'arité de l'opération considérée.

Etablissons tout d'abord un certain nombre de lemmes :

Lemme 2.14

Soit t une opération syntaxique unaire de degré 0 associée à l'opérateur (fidèle) T. Si $\mathcal{L}$ est un cône rationnel (fidèle, fermé par homomorphisme à effacement linéaire) fermé par t et par produit, alors $\mathcal{L}$ est fermé par une opération syntactique binaire t', de degré 0.

Démonstration

Soit t' l'opération binaire définie par

$$t'(L_1,L_2) = t(L_1.L_2)$$

Soit T' l'opérateur (fidèle) associé à t' défini par :

$$T'(\mathcal{L}_1,\mathcal{L}_2) = \mathcal{C}^U(\{t'(L_1,L_2) \mid L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2\})$$

$$(T'(\mathcal{L}_1,\mathcal{L}_2) = \mathcal{C}^{fU}(\{t'(L_1,L_2) \mid L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2\})).$$

Vérifions que t' est de degré 0. Soient donc L_1 et L_2 deux langages strictement disjoints, $\mathcal{L}_1$ et $\mathcal{L}_2$ deux cônes rationnels (fidèles) tels que $t'(L_1,L_2) \in T'(\mathcal{L}_1,\mathcal{L}_2)$. Comme nous avons de façon évidente $T'(\mathcal{L}_1,\mathcal{L}_2) = T(\mathcal{L}_1 \odot \mathcal{L}_2)$, cette condition est équivalente à $t(L_1.L_2) \in T(\mathcal{L}_1 \odot \mathcal{L}_2)$ et donc $L_1.L_2 \in \mathcal{L}_1 \odot \mathcal{L}_2$ puisque t est de degré 0. Comme le produit de deux langages est de degré 0, il en résulte que $L_1 \in \mathcal{L}_1$ ou $L_2 \in \mathcal{L}_2$, i.e. t' est de degré 0.

Soit maintenant $\mathcal{L}$ un cône rationnel (fidèle) tel que $t'(L_1,L_2) = t(L_1.L_2) \in \underline{\text{Rat}}\ \mathcal{L}$ (resp. $\underline{\text{Rat}}_0\ \mathcal{L}$). Comme t est contraint par Rat il en résulte que $L_1L_2 \in \mathcal{L}$ et donc L_1 et L_2 appartiennent à $\mathcal{L}$.

Pour achever la démonstration il nous suffit de remarquer que

puisque $\mathcal{L}$ est fermé par produit et par t, alors $t(L_1.L_2) = t'(L_1,L_2)$ appartient à $\mathcal{L}$ pour tous langages L_1 et L_2 dans $\mathcal{L}$.

∎

Lemme 2.15

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé pour une opération syntactique t n-aire, $n \geq 2$, strictement de degré $p \geq 0$ associée à un opérateur (fidèle) T. Si $\mathcal{L}_1$ et $\mathcal{L}_2$ sont des cônes rationnels fidèles tels que $\mathcal{L} \subseteq \mathcal{L}_1 \vee \mathcal{L}_2$ alors $\mathcal{L}$ est inclus dans $\mathcal{E}(\mathcal{L}_1)$ ou dans $\mathcal{E}(\mathcal{L}_2)$ ($\mathcal{L} \subseteq \mathcal{L}_1$ ou $\mathcal{L} \subseteq \mathcal{L}_2$).

Démonstration

Comme t est strictement de degré p nous pouvons supposer, comme dans [Gr3] que t est une opération n-aire, $n \geq 2$ de degré 0.

Plaçons-nous alors dans le cas d'un opérateur T et d'un cône rationnel $\mathcal{L}$, les autres cas étant semblables. Supposons que $\mathcal{L}$ n'est pas inclus dans $\mathcal{E}(\mathcal{L}_2)$. Soit alors $L_1 \in \mathcal{L} \setminus \mathcal{E}(\mathcal{L}_2)$ et considérons un langage L_2 appartenant à $\mathcal{L}_1 \vee \mathcal{L}_2 \setminus \mathcal{E}(\mathcal{L}_1)$. Nous pouvons supposer que L_1 et L_2 sont strictement disjoints. Si L_2 appartient à $\mathcal{L}$, alors soient $M_1, M_2, \ldots, M_{n-1}$ des copies strictement disjointes de L_2. Le langage $t(L_1, M_1, \ldots, M_{n-1})$ appartient à $\mathcal{L} \subseteq \mathcal{L}_1 \vee \mathcal{L}_2 \subseteq T(\mathcal{E}(\mathcal{L}_2), \mathcal{E}(\mathcal{L}_1)), \ldots, \mathcal{E}(\mathcal{L}_1))$. Comme t est de degré 0, ceci implique $L_1 \in \mathcal{E}(\mathcal{L}_2)$ ou $L_2 \in \mathcal{E}(\mathcal{L}_1)$, ce qui est contradictoire.

∎

Lemme 2.16

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par une opération syntactique strictement de degré $p \geq 0$. Alors

1) $\underline{\text{Rat}}\ \mathcal{L}$ (resp. $\underline{\text{Rat}}_0\ \mathcal{L}$) est indécomposable.

2) $\underline{\text{Rat}}\ \mathcal{L}$ (resp. $\underline{\text{Rat}}_0\ \mathcal{L}$) est principal si et seulement si est principal.

3) Soit $\mathcal{L}' = \{L \in \mathcal{L} \mid \mathcal{C}^U(L) \neq \mathcal{L}\}$

(resp. $\mathcal{L}' = \{L \in \mathcal{L} \mid \mathcal{C}^f(L \cup \{\varepsilon\}) \neq \mathcal{L} \cup \{\varepsilon\}\}$),

alors $\mathcal{L}'$ est un cône rationnel (fidèle) fermé par t et $\underline{\text{Rat}}\ \mathcal{L}' \not\subseteq \text{Ng}(\underline{\text{Rat}}\ \mathcal{L})$ (resp. $\underline{\text{Rat}}_0(\mathcal{L}') \not\subseteq \text{Ng}^f(\underline{\text{Rat}}_0\ \mathcal{L})$).

Démonstration

La démonstration des points 1 et 3 est absolument semblable à celle du lemme 3.2 de [Gr3]. Le point 2 résulte directement du théorème 2.4.

■

On peut alors énoncer les résultats fondamentaux de cette partie :

Théorème 2.13

Soit $\mathcal{L}$ une FAL (fidèle fermée par homomorphisme à effacement linéaire) principale fermée pour un opérateur syntactique (fidèle) n-aire, $n \geq 1$, T. Alors

1) $\text{Ng}(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{C}(L) \neq \mathcal{L}\}$

$(\text{Ng}^f(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{C}^f(L \cup \{\varepsilon\}) \neq \mathcal{L} \cup \{\varepsilon\}\})$.

2) $\mathcal{L}$ est indécomposable.

3) $\text{Ng}(\mathcal{L})$ $(\text{Ng}^f(\mathcal{L}))$ est une FAL (fidèle) fermée par T et est la seule FAL (fidèle) recouverte par $\mathcal{L}$.

4) Si $\mathcal{L}$ est fermé par une opération t' de degré 0 ou indépendante de $\underline{\text{Rat}}$ $(\underline{\text{Rat}}_0)$ alors $\text{Ng}(\mathcal{L})$ $(\text{Ng}^f(\mathcal{L}))$ est fermé par t'.

Démonstration

Si T est un opérateur associé à une opération de degré $p > 0$, alors $\mathcal{L}$ est fermé pour une opération syntactique de degré 0. Si celle-ci est unaire, alors d'après le lemme 2.14, $\mathcal{L}$ est fermée pour une opération binaire de degré 0. Aussi pouvons-nous supposer que t est une opération n-aire, $n \geq 2$, de degré 0. Comme t est indépendante de Rat il résulte du corollaire 2.12 que $Ng(\mathcal{L}) = \{L \in \mathcal{L} \mid \mathcal{C}(L) \neq \mathcal{L}\}$. Le point 2) résulte du lemme 2.16. Le point 3) découle du point 1), du lemme 2.16 et du théorème 2.11. Le point 4) résulte immédiatement des propriétés des opérations de degré 0 et des opérations indépendantes de Rat.

■

Théorème 2.14

Si $\mathcal{L}$ est un cône rationnel (fidèle) qui n'est pas fermé par une opération syntaxique t alors Rat $\mathcal{C}_t^U(\mathcal{L})$ (Rat_0 $\mathcal{C}_t^{fU}(\mathcal{L})$) et $\mathcal{F}_t(\mathcal{L})$ ($\mathcal{F}_t^f \mathcal{L}$)) sont des FAL (fidèles) non principales qui ne recouvrent aucune FAL (fidèle).

Démonstration

Immédiate puisque $\mathcal{C}_t^U(\mathcal{L})$ et $\mathcal{F}_t(\mathcal{L})$ sont non principaux d'après le théorème 2.10.

■

Exemples

1) Dans [Gr3], S. Greibach démontre que le langage $L_= = \{a^n c b^n \mid n \in \mathbb{N}\}$ est indécomposable. Comme $L_=$ est un langage borné, $\mathcal{F}(L_=)$ n'est fermé par aucune opération indépendante de Rat. Ceci montre qu'il existe

des FAL indécomposables qui ne sont fermées pour aucune opération syntactique [Gr3].

Remarquons d'autre part que toute FAL est fermée par produit, i.e. par une opération de degré 0. Le problème reste ouvert de savoir si une FAL indécomposable est nécessairement fermée par une autre opération de degré 0. Nous supposons que la réponse est non en général.

2) Soit $L_1 = \{wcw | w \in \{a,b\}^*\}$. Nous pouvons démontrer par un procédé analogue à celui utilisé pour S_2 dans [Gr3] que L_1 est indécomposable. En effet (cf chapitre 3) si $L_1 = L_1' \cup L_2'$ avec L_1' et L_2' dans $\mathfrak{C}(L_1)$, alors il existe deux langages rationnels R_1 et R_2 tels que

$$L_i' = \{w_i c w_i | w_i \in R_i\},\ i = 1,2 \text{ et}$$

$$R_1 \cup R_2 = \{a,b\}^* \text{ [Rod5]}.$$

Deux cas sont alors possibles :

a) $\exists w_0 \in \{a,b\}^*$ tel que $w_0 \{a,b\}^* \subseteq R_1$, alors le langage

$$L_1'' = L_1' \cap (w_0 \{a,b\}^* c \{a,b\}^* w_0) = w_0 L_1 w_0 \in \mathfrak{C}(L_1')$$

et donc $L_1 \in \mathfrak{C}(L_1')$. Donc $\mathfrak{F}(L_2') \subseteq \mathfrak{F}(L_1) \subseteq \mathfrak{F}(L_1')$ et donc $\mathfrak{F}(L_1')$ et $\mathfrak{F}(L_2')$ ne sont pas incomparables.

b) Sinon pour tout $f \in \{a,b\}^*$, il existe $x \in \{a,b\}^*$ tel que $fx \in R_2$. Comme R_2 est rationnel, il existe un entier n tel que $wx \in R_2$ pour $|x| \leq n$. Soit alors

$$R_3 = \{x \in \{a,b\}^* | |x| \leq n\}.$$

R_3 est fini, si h_x est alors l'homomorphisme défini par $h_x(a) = a$,

$h_x(b) = b$ et $h_x(c) = xcx$, alors

$$L_1 = \bigcup_{x \in R_3} h_x^{-1}(L_2') \cap (\{a,b\}^* \subset \{a,b\}^*), \text{ et donc}$$

$$\mathcal{F}(L_1') \subseteq \mathcal{F}(L_1) \subseteq \mathcal{F}(L_2') \text{ et } L_1' \text{ et } L_2'$$

ne sont pas incomparables.

Nous verrons enfin dans le chapitre 3 que $\mathcal{F}(L_1)$ ne contient pas de FAL non triviale fermée par une opération indépendante de Rat, ce qui nous donne un nouvel exemple de FAL indécomposable non close par une opération syntactique.

3) Dans le cas où $\mathcal{L}$ est un cône rationnel (fidèle) fermé par produit, homomorphisme à effacement linéaire et par une opération syntactique t, par union et par produit on peut établir sans difficulté un théorème analogue au théorème 2.13 :

Théorème 2.14

Soit $\mathcal{L}$ un cône rationnel (fidèle) principal fermé par union, produit, homomorphisme à effacement linéaire et par une opération syntactique t. Alors $\mathcal{L}' = \{L \in \mathcal{L} \mid \mathcal{C}(L) \neq \mathcal{L}\}$ ($\mathcal{L}' = \{L \in \mathcal{L} \mid \mathcal{C}^f(L \cup \{\varepsilon\}) \neq \mathcal{L} \cup \{\varepsilon\}\}$) est un cône rationnel (fidèle) fermé par t et c'est le seul cône rationnel (fidèle) recouvert par $\mathcal{L}$.

Exemples

Dans [BNP,BN] les auteurs définissent la famille $\mathcal{L}_{BNP}$ comme le plus petit cône rationnel fidèle fermé par intersection contenant $S_2 = \{wc\tilde{w} \mid w \in \{a,b\}^*\}$. De façon parallèle dans [BGW] et [BG] est définie la famille $\mathcal{L}_{DUP}$, plus petit cône rationnel fidèle fermé par intersection contenant $L_1 = \{wcw \mid w \in \{a,b\}^*\}$. Ces deux familles sont fermées par homomorphisme à effacement linéaire et l'on démontrera, dans le chapitre suivant, qu'elles sont principales et fermées par chevron.

Il résulte alors du théorème 2.14 que $\mathcal{L}_{BNP}$ et $\mathcal{L}_{DUP}$ recouvrent chacune un unique cône rationnel fidèle fermé par chevron qui n'est autre que l'ensemble des non générateurs fidèles de ces familles. De plus comme $\mathcal{L}_{DUP}$ est une FAL fidèle [BGW] il en résulte que le cône fidèle maximal recouvert par $\mathcal{L}_{DUP}$ est aussi une FAL fidèle.

CHAPITRE 3

A) L'OPERATION CROCHET OUVERT

B) LES OPERATIONS UNAIRES CROCHET, CHEVRON ET CROCHET OUVERT DE DEGRE i

C) LES OPERATIONS UNAIRES CROCHET, CHEVRON ET CROCHET OUVERT DE DEGRE INFINI

A) L'OPERATION CROCHET OUVERT

Nous allons étudier dans cette partie une nouvelle opération unaire appelée crochet ouvert. Cette opération a été introduite essentiellement pour étudier le cône rationnel engendré par le langage $L_1 = \{wcw | w \in \{a,b\}^*\}$ ainsi que la famille des langages vérifiables [Rod1] ("Checking automata Languages" dans [Gr5]). Cette opération associe à tout langage L le langage $CO(L) = \{w\ell w | w \in \{a,b\}^*, \ell \in L, \{a,b\} \cap X_L = \emptyset\}$. La similitude des définitions des opérations crochet [BCN] et crochet ouvert permettait d'espérer des similitudes de propriétés. En fait le seul point commun à ces deux opérations est leur indépendance vis-à-vis de Rat. D'après les chapitres précédents ceci implique notamment que le cône rationnel engendré par L_1 présente de nombreuses analogies structurelles avec celui de $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$: ces deux cônes sont indécomposables et Rat et Rat_0 sont les seuls cônes rationnels fidèles fermés par produit contenus dans $\mathcal{C}(L_1)$ et $\mathcal{C}(S_2)$.

On sait que l'opérateur crochet est idempotent (i.e. $C(\mathcal{L}) = C(C(\mathcal{L}))$), nous montrons par contre que l'opération crochet ouvert est syntactique de degré 0 et régulière. D'autre part en étudiant CO et CO^f l'opérateur et l'opérateur fidèle associés à l'opération crochet ouvert nous démontrons que si L est un langage effaçable (i.e. $\mathcal{C}^f(L) = \mathcal{C}(L)$) alors CO(L) a la même propriété (théorème 3.2). Remarquant qu'il en va de même pour l'opération crochet, ceci nous permet d'exhiber une infinité de langages (algébriques ou non) qui sont effaçables.

Définition 3.1

Le crochet ouvert d'un langage L par A appartenant à Rat_2, noté CO(L,A) est le langage

$$CO(L,A) = \{awb | (a,b) \in A, w \in L\}.$$

Une famille de langages $\mathcal{L}$ est fermée par crochet ouvert si et seulement si pour tout A dans Rat_2 et tout L dans $\mathcal{L}$, le langage CO(L,A) appartient à $\mathcal{L}$.

Nous désignerons par $CO(\mathcal{L})$ (resp. $CO^f(\mathcal{L})$) le plus petit cône rationnel (resp. fidèle) fermé par union et contenant la famille $\mathcal{L}'$ des crochets ouverts des langages L de $\mathcal{L}$ par $A \in Rat_2$:

$$\mathcal{L}' = \{CO(L,A)/L \in \mathcal{L},\ A \in Rat_2\}.$$

Dans le cas où $\mathcal{L}$ est un cône rationnel (fidèle) nous pouvons énoncer :

Théorème 3.1 [Rod5]

Si $\mathcal{L}$ est un cône rationnel (fidèle) alors $CO(\mathcal{L})$ ($CO^f(\mathcal{L})$) est identique à la fermeture par union de la famille

$$\mathcal{L}' = \{CO(L,A) \mid L \in \mathcal{L},\ A \in Rat_2\}.$$

Démonstration

Il est clair que $CO(\mathcal{L})$ contient la fermeture par union de la famille $\mathcal{L}'$. Il nous suffit donc de vérifier que la fermeture par union de $\mathcal{L}'$ est un cône rationnel.

D'après le théorème de caractérisaiton des transductions rationnelles de Nivat [Ni1], on sait que pour toute transduction rationnelle τ de X dans Y^* il existe un monoïde libre Z^*, un langage rationnel $K \subseteq Z^*$ et deux homomorphismes alphabétiques φ et ψ de Z^* dans X^* et Y^* respectivement, tels que $\forall f \in X^*$, $\tau(f) = \psi(\varphi^{-1}f \cap K)$.

Soit donc $L' \in \mathcal{L}'$, $A \in Rat_2$ et $L \in \mathcal{L}$ tels que $L' = CO(L,A)$.

a) Soit φ un homomorphisme alphabétique de U^* dans T^* avec $L' \subseteq T^*$. Comme φ est alphabétique, on a :

$$\varphi^{-1}(L') = CO(\varphi^{-1}L, \varphi^{-1}A).$$

$\varphi^{-1}L$ appatient au cône rationnel $\mathcal{L}$ et $\varphi^{-1}A$ appartient à Rat_2 [BCN].

Donc $\varphi^{-1}(L') \in \mathcal{L}'$.

b) Soit φ un homomorphisme de monoïdes libres. Alors $\varphi(L') = CO(\varphi L, \varphi A)$ appartient à $\mathcal{L}'$ puisque $\varphi L \in \mathcal{L}$ et $\varphi A \in Rat_2$ d'après la caractérisation de [Ni1].

c) Soit $K \in Rat(X^*)$ reconnu par un automate fini ayant Q pour ensemble d'états, q_0 pour état initial, F pour états finals et δ pour fonction de transition.

Pour tout q et q' $\in Q$, soient

$$K_q = \{w \in X^* | \delta(q_0,w) = q\}$$

$$K_{qq'} = \{w \in X^* | \delta(q,w) = q'\}$$

$$K'_q = \{w \in X^* | \delta(q,w) \in F\}$$

ces ensembles sont rationnels et $K = \bigcup_{qq'} K_q K_{qq'} K'_q$.

Soit $L' = CO(L,A) \subseteq X^*$. On a clairement

$$L' \cap K = \bigcup_{qq' \in Q} CO(L \cap K_{qq'}, A \cap (K_q \times K'_{q'}))$$

or $L \cap K_{qq'}$ appartient au cône rationnel $\mathcal{L}$ et si $A = \{(\varphi w, \psi w) | w \in R\}$ alors $A \cap (K_q \times K'_{q'}) = \{(\varphi w, \psi w) | w \in R \cap \varphi^{-1}(K_q) \cap \psi^{-1}(K'_{q'})\}$ et donc $A \cap (K_q \times K_{q'})$ appartient à Rat_2.

Chacun des crochets ouverts $CO(L \cap K_{qq'}, A \cap (K_q \times K'_{q'}))$ appartient donc à $\mathcal{L}'$ et leur union finie à la fermeture par union de $\mathcal{L}'$. D'où le résultat.

Dans le cas où le $\mathcal{L}$ est un cône rationnel fidèle, la preuve est

identique en supposant au point b) que φ est un homomorphisme non effaçant.

■

Il est à noter que si $\mathcal{L}$ est un cône rationnel alors le théorème précédent implique :

Corollaire 3.1

Si $\mathcal{L}$ est un cône rationnel, alors $CO(\mathcal{L}) = CO^f(\mathcal{L})$.

Il est clair que l'opérateur crochet ouvert CO est simple i.e. $\mathcal{L} \subseteq CO(\mathcal{L})$, pour tout cône rationnel $\mathcal{L}$. Le crochet ouvert d'un langage L par A appartenant à Rat_2 n'est pas une opération au sens où nous l'avons défini dans le chapitre précédent. Montrons maintenant que l'opérateur crochet ouvert est associé à l'opération unaire crochet ouvert du langage L, définie par :

Définition 3.2

Soit $L \subseteq X^*$ un langage et $\{a,b\}$ un alphabet disjoint de X. Le crochet ouvert de L est le langage

$$CO(L) = \{w\ell w \mid w \in \{a,b\}^*, \ \ell \in L\}.$$

Il est à noter que CO(L) est défini à un isomorphisme alphabétique près.

On a alors

Lemme 3.1

L'opération crochet ouvert est régulière.

Démonstration

Celle-ci est évidente car si L domine rationnellement L' alors on obtient CO(L') à partir de CO(L) par la transduction rationnelle qui recopie le préfixe $w \in \{a,b\}^*$ de tout mot inclus dans $\{a,b\}^* X_L^* \{a,b\}^*$, fonctionne alors comme la transduction rationnelle τ_1 qui vérifie $L' = \tau_1(L)$ et recopie alors le suffixe inclus dans $\{a,b\}^*$.

■

Lemme 3.2 [Rod5]

Soit L un langage, alors

$$CO(\mathcal{C}(L)) = \mathcal{C}(CO(L)) \quad (CO^f(\mathcal{C}^f(L)) = \mathcal{C}^f(CO(L))).$$

Démonstration

Soit $L_0 \in CO(\mathcal{C}(L))$, nous allons montrer que L_0 est image de CO(L) dans une transduction rationnelle. Comme CO(L) appartient à $CO(\mathcal{C}(L))$ le résultat en découlera.

D'après le théorème 3.1, il existe $n \in \mathbb{N}$, n ensembles $A_j \in Rat_2$ et et n langages $M_j \in \mathcal{C}(L)$ tels que $L_0 = \bigcup_{j=1}^{n} CO(M_j,A_j)$. Comme $\mathcal{C}(L)$ est un cône rationnel principal il existe n transductions rationnelles τ_j, $1 \leq j \leq n$, telles que $M_j = \tau_j(L_j)$, $1 \leq j \leq n$, où $L_1,\ldots,L_n$ sont des copies de L sur des alphabets $X_{L_1},\ldots,X_{L_n}$ disjoints deux à deux. On peut supposer que ε n'appartient pas à L_j, $1 \leq j \leq n$.

D'autre part pour tout $j \in [1,n]$, il existe n langages rationnels

locaux K_j et 2n homomorphismes alphabétiques φ_j et ψ_j, $1 \leq j \leq n$, tels que

$$A_j = \{(\varphi_j w, \psi_j w) | w \in K_j\} \qquad [Ni1]$$

On peut supposer les alphabets X_{K_j} et X_{L_j} disjoints deux à deux.

Soit alors X un alphabet d'au moins deux lettres, disjoint des précédents. Soit pour tout j, $1 \leq j \leq n$, $L_{0_j} = CO(L_j)$. Il est facile de vérifier, étant donné la disjonction des alphabets qu'il existe des transductions rationnelles τ'_j telles que

$$L'_j = CO(L_j, \{(w,w) | w \in K_j\}) = \tau'_j(L_{0_j})$$

Soit alors pour tout j, $1 \leq j \leq n$, τ''_j la transduction rationnelle qui rejette les mots n'appartenant pas à $X^*_{K_j} X^*_{L_j} X^*_{K_j}$ et fonctionne comme φ_j sur le préfixe des mots inclus dans $X^*_{K_j}$, comme τ_j sur le facteur inclus dans $X^*_{L_j}$ et comme ψ_j sur le suffixe dans $X^*_{K_j}$. τ''_j se construit facilement étant donné la disjonction des alphabets. On obtient alors :

$$CO(M_j, A_j) = \tau''_j\ CO(L_j, \{(w,w) | w \in K_j\})$$

et par conséquent

$$L_0 = \bigcup_{j=1}^{n} \tau''_j \circ \tau'_j\ CO(L_j)$$

et donc L_0 est bien image par transduction rationnelle de CO(L), puisque le cône rationnel $\mathfrak{C}(CO(L))$ est fermé par union puisque principal.

Considérons maintenant le cas de $\mathfrak{C}^f(CO(L))$. Le langage CO(L) appartient à $CO^f(\mathfrak{C}^f(L))$ qui est un cône rationnel fidèle donc

$\mathcal{E}^f(CO(L)) \subseteq CO^f(\mathcal{E}^f(L))$. Réciproquement si on considère $L_0 \in CO^f(\mathcal{E}^f(L))$ alors d'après le théorème 3.1 il existe comme précédemment des langages $M_j \in \mathcal{E}^f(L)$ et des ensembles $A_j \in Rat_2$ tels que $L_0 = \bigcup_{j=1}^{n} CO(M_j,A_j)$. Si $A_j = \{(\varphi_j w,\psi_j w) | w \in K_j\}$ et φ_j et ψ_j sont non effaçants ou à effacement borné (i.e. pour tout mot $h \in K_j$, le nombre de lettres consécutives d'image vide par φ_j et ψ_j est borné) alors la preuve est semblable à celle qui précède.

Si tel n'est pas le cas, supposons alors que φ_j peut effacer plus de N lettres consécutives de mots $h \in K_j$, où N est l'entier associé à K_j par le lemme de l'étoile. Alors si $xyz \in K_j$ avec $\varphi_j(y) = \varepsilon$ il en résulte que $y = \alpha\beta\gamma$ avec $|\alpha|$ et $|\gamma|$ inférieurs à N tels que $x\alpha\beta^*\gamma z \subseteq K_j$. L'effacement de α et γ se fait par un homomorphisme à effacement borné sur K_j. Soit alors θ_j l'homomorphisme de $X^*_{K_j}$ dans lui-même défini par $\theta_j(a) = a$ si $\varphi_j(a) \neq \varepsilon$ et $\varphi_j(a) = \varepsilon$ sinon. Alors $A_j = \{(w,\lambda_j(w)) | w \in \theta(K_j)\}$ où φ_j est la transduction rationnelle fidèle définie par

$$\lambda_j(f) = \psi_j(\theta_j^{-1}(f) \cap K_j)$$

Comme nous pouvons supposer que pour tout $a \in X_{K_j}$, $\varphi_j(a)$ et $\psi_j(a)$ ne sont pas simultanément égaux à ε, on en déduit une transduction rationnelle fidèle τ telle que $L_j = \tau(CO(L))$ par un procédé analogue à celui utilisé précédemment.

■

Corollaire 3.2 [Rod5]

L'opérateur crochet ouvert (fidèle) est associé à l'opération crochet ouvert, i.e. pour tout cône rationnel (fidèle) $\mathcal{L}$

$$CO(\mathcal{L}) = \mathcal{E}^{U}(\{CO(L) \mid L \in \mathcal{L}\}) \quad (CO^{f}(\mathcal{L}) = \mathcal{E}^{fU}(\{CO(L) \mid L \in \mathcal{L}\}).$$

Corollaire 3.3 [Rod5]

Si $\mathcal{L}$ est un cône rationnel (fidèle) de générateur L, alors $CO(\mathcal{L})$ $(CO^{f}(\mathcal{L}))$ est un cône rationnel (fidèle) de générateur CO(L).

Corollaire 3.4 [Rod5]

Un cône rationnel (fidèle) principal $\mathcal{L}$ est fermé par crochet ouvert si et seulement si pour tout générateur L de $\mathcal{L}$, le langage CO(L) appartient à $\mathcal{L}$.

Les corollaires 3.1 et 3.3 nous permettent alors d'exhiber une infinité de langages pour lesquels le cône rationnel fidèle engendré est fermé par homomorphisme. Dans [Lat1] M. Latteux appelle langage effaçable tout langage ayant cette propriété.

De façon plus générale, les cônes rationnels tels que pour tout système générateur $\mathcal{S}$ on a $\mathcal{E}^{f}(\mathcal{S}) = \mathcal{E}(\mathcal{S})$ ont été étudiés dans [G.G.G.] Suivant la terminologie de M. Latteux nous définissons :

Définition 3.3

Un langage L est effaçable si $\mathcal{E}^{f}(L) = \mathcal{E}(L)$. Un cône rationnel $\mathcal{L}$ est effaçable si pour tout système générateur $\mathcal{S}$ on a $\mathcal{L} = \mathcal{E}(\mathcal{S}) = \mathcal{E}^{f}(\mathcal{S})$.

On peut alors énoncer :

Corollaire 3.5 [Rod5]

Soit $\mathcal{L}$ un cône rationnel (fidèle) principal de générateur L. Si X est un alphabet d'au moins deux lettres disjoint de X_L et c un nouveau symbole, alors pour tout $i \geq 1$, $CO^i(\mathcal{L})$ (resp. $CO^{fi}(\mathcal{L})$) est un cône rationnel (fidèle) principal de générateur

$$A_i = \{w_1cw_2c\ldots w_ic\ell cw_i\ldots cw_2cw_1 | w_p \in X^*,\ 1 \leq p \leq i,\ \ell \in L\}.$$

De plus si $\mathcal{L}$ est un cône rationnel effaçable alors $CO^i(\mathcal{L})$ est effaçable, i.e. $\mathcal{E}^f(A_i) = \mathcal{E}(A_i)$.

Ce corollaire nous permet donc d'énoncer plus généralement

Théorème 3.2

Pour tout langage effaçable L, le langage CO(L) est effaçable.

Comme Rat est un cône rationnel effaçable nous en déduisons immédiatement, si nous introduisons :

Définition 3.4

Pour tout $i \geq 1$, nous désignerons par L_i le langage

$$L_i = \{w_1cw_2c\ldots cw_idw_ic\ldots cw_2cw_1 | w_p \in \{a,b\}^*,\ 1 \leq p \leq i\}.$$

Notations

Nous écrirons $CO^i(\mathcal{L})$ pour $CO(\ldots(CO(\mathcal{L}))\ldots)$, i.e. l'application i fois répétée de l'opérateur CO sur la famille $\mathcal{L}$. Nous désignerons de même par $CO^{fi}(\mathcal{L})$ la famille $CO^f(\ldots(CO^f(\mathcal{L})))$. On a alors :

Corollaire 3.6

Pour tout $i \geq 1$, le langage L_i est effaçable, i.e.
$CO^i(Rat) = CO^{fi}(Rat) = \mathcal{E}(L_i) = \mathcal{E}^f(L_i)$.

Notons que la preuve du lemme 3.2 s'adapte facilement au cas de l'opérateur crochet (fidèle). Ceci nous permet donc d'énoncer les deux résultats suivants, dont le premier est, semble-t-il, connu des auteurs de [BCN] dans le cas des cônes rationnels :

Lemme 3.3

Si $\mathcal{L}$ un cône rationnel (fidèle) alors

$$C(\mathcal{L}) = \mathcal{E}^U(\{w\ell\tilde{w} \mid \ell \in L, w \in \{a,b\}^*, \{a,b\} \cap X_L = \emptyset\}, L \in \mathcal{L})$$

$$(C^f(\mathcal{L}) = \mathcal{E}^{fU}(\{w\ell\tilde{w} \mid \ell \in L, w \in \{a,b\}^*, \{a,b\} \cap X_L = \emptyset\}, L \in \mathcal{L})).$$

Corollaire 3.7

Pour tout langage effaçable L, le langage C(L) est effaçable, i.e.
$\mathcal{E}^f(C(L)) = \mathcal{E}(C(L))$.

Comme Rat est effaçable, il en résulte

Corollaire 3.8

$S_2 = \{wc\tilde{w} \mid w \in \{a,b\}^*\}$ est effaçable, i.e. $\mathcal{E}^f(S_2) = \mathcal{E}(S_2)$.

Remarque

Les théorèmes précédents nous permettent donc d'exhiber, en général, une infinité de langages effaçables à partir de tout langage effaçable L donné. En effet si L n'est pas rationnellement équivalent à C(L), ni à CO(L), alors C(L) et $CO^i(L)$, $i \geq 1$ sont également effaçables.

Ainsi M. Latteux à établi dans [Lat1] que tout langage algébrique L inclus dans a^*b^* est effaçable. Nos résultats permettent de conclure que les langages algébriques correspondants C(L) sont également effaçables, i.e. d'exhiber une infinité de langages algébriques non bornés ayant cette propriété. Il en va de même pour les langages vérifiables (cf chapitre 7 de ce travail) associés $CO^i(L)$, pour tout $i \geq 1$.

Nous allons maintenant démontrer que $CO(CO^f)$ est un opérateur syntactique (fidèle). Pour cela rappelons tout d'abord un résultat de M. Latteux [Lat4] :

Théorème 3.3 [Lat4]

L'opération chevron est indépendante de l'opérateur crochet ouvert (fidèle), i.e. pour tout langage L et tout cône rationnel $\mathcal{L}$ la condition $Ch(L) \in CO(\mathcal{L})$ implique $L \in \mathcal{L}$ ($Ch(L) \in CO^f(\mathcal{L}) \Longrightarrow L \in \mathcal{L}$).

Il en résulte immédiatement :

Théorème 3.4 [Rod5]

CO et CO^f sont des opérateurs de degré 0.

Démonstration

Si $\varepsilon \notin L$ alors la condition $CO(L) \in CO(\mathcal{L})$ implique trivialement $Ch(L) \in CO(\mathcal{L})$ et donc $L \in \mathcal{L}$ d'après le théorème précédent.

Si $\varepsilon \in L$ et $CO(L) \in CO(\mathcal{L})$, alors
$CO(L-\{\varepsilon\}) = CO(L) \cap X^*X_L X_L^*X^* \in CO(\mathcal{L})$ et de même que précédemment $L-\{\varepsilon\} \in \mathcal{L}$. Comme $\mathcal{L}$ est un cône rationnel fermé par union, il en résulte que $L \in \mathcal{L}$.

La preuve est identique dans le cas de CO^f.

■

Du lemme 2.11 et du théorème 2.8, il en résulte alors immédiatement :

Corollaire 3.9 [Rod5]

Soit $\mathcal{L}$ un cône rationnel fermé par union et non clos par crochet ouvert, alors $CO^i(\mathcal{L}) \subsetneq CO^{i+1}(\mathcal{L})$, $\forall i \geq 1$, $(CO^{fi}(\mathcal{L}) \subsetneq CO^{f(i+1)}(\mathcal{L}))$.

Corollaire 3.10 [Rod5]

Soit $\mathcal{L}$ un cône rationnel fermé par union et non clos par crochet ouvert (fidèle), alors $\mathcal{E}^{U}_{CO}(\mathcal{L})$ (resp. $\mathcal{E}^{fU}_{CO}(\mathcal{L})$) le plus petit cône rationnel fermé par union et crochet ouvert (resp. fidèle) contenant $\mathcal{L}$ est non principal.

Corollaire 3.11 [Rod5]

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par union, L un langage et X un alphabet d'au moins deux lettres disjoint de X_L. Si

$$C(L) = \{w\ell\tilde{w} \mid \ell \in L, w \in X^*\} \in \mathcal{E}^{U}_{CO}(\mathcal{L})$$

(resp. $\mathcal{E}^{fU}_{CO}(\mathcal{L})$) alors $C(L) \in \mathcal{L}$.

Démonstration

Il suffit de remarquer qu'il existe une transduction rationnelle fidèle τ telle que $Ch^p(L) = \tau(C(L))$, où Ch^p est l'application p fois répétée de l'opération chevron sur le langage L, $\forall p \geq 1$. Si donc $C(L) \in \mathcal{E}^{U}_{CO}(\mathcal{L})$, il existe $p \in \mathbb{N}$ tel que $C(L) \in CO^p(\mathcal{L})$, donc $Ch^p(L) \in CO^p(\mathcal{L})$. Le résultat découle alors de l'application, p fois répétée, du théorème 3.3.

∎

Il en résulte alors immédiatement :

Corollaire 3.12

Soit $\mathcal{L}$ un cône rationnel fermé par union et $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$. Alors $S_2 \in \mathcal{E}^{U}_{CO}(\mathcal{L})$ implique $S_2 \in \mathcal{L}$.

En particulier nous en déduisons le résultat suivant, établi de façon tout à fait différente par Klingenstein [Klin] :

Théorème 3.5 [Rod5]

$S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et $L_1 = \{wcw | w \in \{a,b\}^*\}$ sont rationnellement incomparables.

Comme le chevron est indépendant de Rat [BoNi1] il en résulte :

Corollaire 3.13 [Rod5]

L'opération crochet ouvert est indépendante de Rat (Rat_0).

On a donc :

Théorème 3.6

L'opération crochet ouvert est une opération syntaxique de degré 0.

Nous allons maintenant montrer que le produit de deux langages est indépendant de l'opérateur $CO(CO^f)$.

Théorème 3.7 [Rod5]

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par union, L_1 et L_2 deux langages sur des alphabets disjoints X_1 et X_2. La condition $L_1L_2 \in CO(\mathcal{L})$ $(CO^f(\mathcal{L}))$ implique $L_1 \in Rat$ ou $L_2 \in \mathcal{L}$.

Démonstration

Si $L_1L_2 \in CO(\mathcal{L})$, alors d'après le théorème 3.1 il existe un entier n, n ensembles $A_j \in Rat_2$ et n langages $M_j \in \mathcal{L}$, $1 \leq j \leq n$, tels que

$$L_1L_2 = \bigcup_{j=1}^{n} CO(M_j,A_j)$$

Nous supposerons que A_j et M_j sont non vides, $1 \leq j \leq n$. Pour tout $w_1 \in L_1$ et tout $w_2 \in L_2$, il existe donc j, $1 \leq j \leq n$, $(a,b) \in A_j$ et $\ell \in M_j$ tels que $w_1w_2 = a\ell b$. Deux cas se présentent :

a) Pour tout $w_1 \in L_1$, il existe $w_2 \in L_2$, $1 \leq j \leq n$, $(a,b) \in A_j$ $\ell \in M_j$ tels que $w_1w_2 = a\ell b$ et w_1 est un facteur gauche propre de a.

Soit alors

$$L = \{w \in X_1^* \mid \exists w' \in X_2X_2^*, b \in X_2^*, 1 \leq j \leq n : (ww',b) \in A_j\}.$$

Montrons que $L_1 = L$.

Les conditions de ce cas impliquent $L_1 \subseteq L$. Réciproquement, soit $w \in X_1^*$ tel qu'il existe $1 \leq j \leq n$, $w' \in X_2^*X_2$, $b \in X_2^*$ vérifiant $(ww',b) \in A_j$. Le langage M_j étant vide, il existe $\ell \in M_j$ et le mot $ww'\ell b \in CO(M_j,A_j) \subseteq L_1L_2$. La disjonction des alphabets implique alors $w \in L_1$. Soit alors $\pi_1(A_j) = \{a \mid \exists b : (a,b) \in A_j\}$. Il est clair que

$$L_1 = \{w \in X_1^* \mid \exists w' \in X_2X_2^* : ww' \in \bigcup_{j=1}^{n} \pi_1(A_j)\}.$$

Or $\bigcup_{j=1}^{n} \pi_1(A_j)$ est un langage rationnel et L_1 est image de ce dernier dans une transduction rationnelle évidente. Donc L_1 est un langage rationnel.

b) Il existe $w_1 \in L_1$ tel que pour tout $w_2 \in L_2$, $1 \leq j \leq n$, $(a,b) \in A_j$ et $\ell \in M_j$, la condition $w_1 w_2 = a\ell b$ implique que a est facteur gauche de w_1.

Vérifions que L_2 est égal au langage

$$L_2' = \{w'b \in X_2^* \mid \exists a,w \in X_1^*,\ 1 \leq j \leq n : aw = w_1, ww' \in M_j, (a,b) \in A_j\}.$$

Les conditions de ce cas impliquent $L_2 \subseteq L_2'$. Réciproquement soit $w'b \in X_2^*$ tel qu'il existe a, $w \in X_1^*$, $1 \leq j \leq n$, vérifiant $aw = w_1$, $ww' \in M_j$ et $(a,b) \in A_j$. On a donc $aww'b \in CO(M_j,A_j) \subseteq L_1L_2$ et la disjonction des alphabets implique $w'b \in L_2$.

Pour tout w_1' facteur droit de w_1 et tout $1 \leq j \leq n$, soient

$$P_j(w_1') = \{w_1'' \mid w_1'w_1'' \in M_j\}$$

$$Q_j(w_1') = \{b \mid (a,b) \in A_j,\ aw_1' = w_1\}$$

On a alors

$$L_2 = \bigcup_{w_1'} \{P_j(w_1')\ Q_j(w_1') \mid 1 \leq j \leq n\}.$$

Or $P_j(w_1')$ est image de M_j par la transduction rationnelle qui rejette les mots ne commençant pas par w_1' et efface w_1' au début des autres. De plus $Q_j(w_1')$ est un langage rationnel image de a tel que $aw_1' = w_1$ par la transduction associée à A_j. Il en résulte donc que $P_j(w_1')Q_j(w_1')$ est image de M_j dans une transduction rationnelle $\tau_{j,w_1'}$. Nous pouvons donc écrire

$$L_2 = \bigcup_{w_1'} \{\tau_{j,w_1'}(M_j) \mid 1 \leq j \leq n\}$$

i.e. L_2 appartient à $\mathcal{L}$ puisque $\mathcal{L}$ est fermé par union.

Dans le cas de l'opérateur CO^f la preuve est identique et ne nécessite que les précautions habituelles pour montrer que M_j est image de $P_j(w_1')$ dans une transduction rationnelle fidèle.

■

On a alors

Corollaire 3.14 [Rod5]

Soit $\mathcal{L}$ un cône rationnel fermé par union. La condition $\mathcal{L} \subsetneq CO(\mathcal{L})$ implique que le plus petit cône rationnel fermé par union, produit et crochet ouvert contenant $\mathcal{L}$ est non principal.

Le théorème 3.7 nous permet d'établir pour le langage $L_1 = \{wcw \mid w \in \{a,b\}^*\}$ des propriétés analogues à celles vues au chapitre 1, Section E pour les langages linéaires. En effet nous pouvons énoncer :

Théorème 3.8

Soit $\mathcal{L}$ un cône rationnel (fidèle) non fermé par crochet ouvert, alors :

1) Tout cône rationnel (fidèle) fermé par produit contenu dans $CO^i(\mathcal{L})$ ($CO^{fi}(\mathcal{L})$) est contenu dans $\mathcal{L}$.

2) Toute FAL (fidèle) contenue dans $\mathcal{E}^U_{CO}(\mathcal{L})$ ($\mathcal{E}^{fU}_{CO^f}(\mathcal{L})$) est contenue dans $\mathcal{L}$.

3) $\mathcal{E}^{U}_{CO}(\mathcal{L})$ ($\mathcal{E}^{fU}_{CO^f}(\mathcal{L})$) n'est pas une FAL fidèle et n'est pas un cône rationnel (fidèle) principal.

4) $\underline{Rat}\ \mathcal{E}^{U}_{CO}(\mathcal{L})$ ($\underline{Rat}_0\ \mathcal{E}^{fU}_{CO^f}(\mathcal{L})$) n'est pas fermé par crochet ouvert, est indécomposable et n'est pas une FAL (fidèle) principale.

5) Tout cône rationnel fermée par produit et une opération indépendante de $\underline{Rat}$ contenue dans $\underline{Rat}\ CO(\mathcal{L})$ est contenu dans $\mathcal{L}$.

Démonstration

1) Soit L un langage appartenant à $CO^i(\mathcal{L})\setminus\mathcal{L}$, et c un nouveau symbole. Si $(Lc)^i$ appartient à $CO^i(\mathcal{L})$, alors d'après le théorème 3.7 $(Lc)^{i-1}$ appartient à $CO^{i-1}(\mathcal{L})$ puisque L n'appartient pas à Rat. Une récurrence descendante triviale conduit alors à une contradiction, d'où le résultat.

2) Si $(Lc)^+$ appartient à $\mathcal{E}^{U}_{CO}(\mathcal{L})$ alors il existe i tel que $(Lc)^+ \in CO^i(\mathcal{L})$ et le résultat découle du 1).

3) Du point 2) il résulte que $\underline{Rat}_0\ CO^i(\mathcal{L})$, $i \geq 1$, n'est pas inclus dans $\mathcal{E}^{U}_{CO}(\mathcal{L})$, donc $\mathcal{E}^{U}_{CO}(\mathcal{L})$ n'est pas une FAL fidèle. D'après le théorème 3.6 et le corollaire 2.2 , $\mathcal{E}^{U}_{CO}(\mathcal{L})$ est non principal.

4) Résulte directement des théorèmes 2.4 et 2.5.

5) Si $\mathcal{L}_1$ est un cône rationnel fermé par produit et par une opération indépendante de $\underline{Rat}$ contenue dans $\underline{Rat}\ CO(\mathcal{L})$ alors $\mathcal{L}_1 \subseteq CO(\mathcal{L})$ ce qui contredit le point 1).

Les preuves sont identiques dans le cas de CO^f.

■

On en déduit immédiatement :

Corollaire 3.15

1) Rat et Rat_0 sont les seuls cônes rationnels fidèles fermés par produit contenus dans le cône rationnel engendré par $L_1 = \{wcw | w \in \{a,b\}^*\}$.

2) Rat et Rat_0 sont les seuls cônes rationnels fidèles fermés par produit et par une opération indépendante de Rat (Rat_0) inclus dans la FAL engendrée par L_1.

B) LES OPERATIONS UNAIRES CHEVRON, CROCHET ET CROCHET OUVERT DE DEGRE i

Nous introduisons, dans cette partie, trois familles d'opérations qui généralisent les opérations chevron, crochet et crochet ouvert considérées précédemment. Ces opérations, notées respectivement Ch_i, C_i et CO_i pour i appartenant à $\mathbb{N}$, sont appelées chevron, crochet et crochet ouvert de degré i. L'intérêt de cette généralisation réside en fait dans la caractérisation de deux familles de langages, $\mathcal{L}_{BNP}$ et $\mathcal{L}_{DUP}$ [BNP,BN,BG] par les opérations C_∞ et CO_∞, limites des suites $\{C_i\}_{i \in \mathbb{N}}$ et $\{CO_i\}_{i \in \mathbb{N}}$ que nous considèrerons dans la section suivante.

Après avoir donné une caractérisation des familles $C_i(\mathcal{L})$ et $CO_i(\mathcal{L})$ analogue au théorème 3.1, dans le cas où $\mathcal{L}$ est un cône rationnel, nous concluons tout d'abord que les opérations C_i et CO_i préservent les langages effaçables (i.e. si $\mathcal{E}^f(L) = \mathcal{E}(L)$ alors $\mathcal{E}^f(C_i(L)) = \mathcal{E}(C_i(L))$). Nous montrons alors que les opérations Ch_i, C_i et CO_i, $i \in \mathbb{N}$, sont indépendantes de Rat et que le produit de deux langages est indépendant des opérateurs associés à ces opérations.

Nous avons vu que les opérations chevron et crochet ouvert étaient de degré 0 tandis que l'opérateur crochet était idempotent : contrairement à toute attente, nous montrons que les opérations chevron, crochet et crochet ouvert de degré i pour $i \geq 2$ sont toutes de degré 0. Ceci nous permet alors de conclure que les plus petits cônes rationnels fermés pour chacune de ces opérations est non principal (corollaire 3.18).

D'une condition nécessaire et suffisante pour que $C_i(\mathcal{L}) = C(\mathcal{L})$ (resp. $CO_i(\mathcal{L}) = CO(\mathcal{L})$) pour tout $i \geq 1$ nous déduisons enfin que les cônes rationnels engendrés par $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et $L_1 = \{wcw | w \in \{a,b\}^*\}$ sont respectivement égaux à $C_i(Rat)$ et $CO_i(Rat)$ pour tout $i \geq 1$.

Définition 3.5

Soit $L \subseteq X_L^*$ un langage et $X = \{a,b\}$ tel que $X_L \cap X = \emptyset$. Nous appellerons pour tout $i \geq 1$:

- Crochet de degré i de L, le langage

$$C_i(L) = \{a_1\ell_1a_2\ell_2\ldots a_i\ell_i\tilde{b}_i\ell_{i+1}\ldots\tilde{b}_2\ell_{2i-1}\tilde{b}_1 \mid \ell_1\ell_2\ldots\ell_{2i-1} \in L, a_j,b_j \in X^*,$$
$$1 \leq j \leq i,\ a_1a_2\ldots a_i = b_1b_2\ldots b_i \in X^*\}.$$

- Chevron de degré i de L, le langage

$$Ch_i(L) = \{a^{n_1}\ell_1a^{n_2}\ell_2\ldots a^{n_i}\ell_ib^{m_1}\ell_{i+1}\ldots b^{m_{i-1}}\ell_{2i-1}b^{m_i} \mid \ell_1\ell_2\ldots\ell_{2i-1} \in L,$$
$$n_j \text{ et } m_j \in \mathbb{N},\ 1 \leq j \leq i,\ \sum_{j=1}^{i} n_j = \sum_{j=1}^{i} m_j\}.$$

- Crochet ouvert de degré i de L, le langage

$$CO_i(L) = \{a_1\ell_1\ldots a_i\ell_ib_1\ell_{i+1}\ldots b_{i-1}\ell_{2i-1}b_i \mid \ell_1\ell_2\ldots\ell_{2i-1} \in L,\ a_j,b_j \in X^*,$$
$$1 \leq j \leq i,\ a_1a_2\ldots a_i = b_1b_2\ldots b_i\}.$$

Notons que les langages $C_i(L)$, $Ch_i(L)$ et $CO_i(L)$ sont définis à un isomorphisme alphabétique près. Il est clair que pour $i = 1$ nous retrouvons les opérations crochet, chevron et crochet ouvert étudiées précédemment.

Définition 3.6

Nous désignerons également par C_i, Ch_i et CO_i les opérateurs associés aux opérations crochet, chevron et crochet ouvert de degré i, i.e. pour toute famille de langages $\mathcal{L}$:

$$C_i(\mathcal{L}) = \mathcal{E}^U(\{C_i(L) | L \in \mathcal{L}\})$$

$$Ch_i(\mathcal{L}) = \mathcal{E}^U(\{Ch_i(L) | L \in \mathcal{L}\})$$

$$CO_i(\mathcal{L}) = \mathcal{E}^U(\{CO_i(L) | L \in \mathcal{L}\})$$

Nous considèrerons enfin C_i^f, Ch_i^f et CO_i^f les opérateurs fidèles associés définis par :

$$C_i^f(\mathcal{L}) = \mathcal{E}^{fU}(\{C_i(L) | L \in \mathcal{L}\})$$

$$Ch_i^f(\mathcal{L}) = \mathcal{E}^{fU}(\{Ch_i(L) | L \in \mathcal{L}\})$$

$$CO_i^f(\mathcal{L}) = \mathcal{E}^{fU}(\{CO_i(L) | L \in \mathcal{L}\}).$$

Nous pouvons alors énoncer :

Lemme 3.4

Pour tout $i \geq 1$, les opérations C_i, Ch_i et CO_i sont régulières.

Démonstration

Si $L' = \tau(L)$ pour quelque transduction rationnelle τ alors soit $X = \{a,b\}$ et $X' = \{c,d\}$ de nouveaux alphabets tels que $X \cap X' = \emptyset$, $X \cap X_L = \emptyset$, $X' \cap X_{L'} = \emptyset$. Etant donné la disjonction des alphabets on construit alors sans difficulté une transduction rationnelle τ' telle que $\tau'(C_i(L)) = C_i(L')$. Il en va de même pour Ch_i et CO_i.

■

Corollaire 3.16

Si $\mathcal{L}$ est un cône rationnel (fidèle) de générateur L, alors

$C_i(\mathcal{L})$, $Ch_i(\mathcal{L})$ et $CO_i(\mathcal{L})$ (resp. $C_i^f(\mathcal{L})$, $Ch_i^f(\mathcal{L})$ et $CO_i^f(\mathcal{L})$) sont des cônes rationnels (fidèles) principaux de générateurs respectifs $C_i(L)$, $Ch_i(L)$ et $CO_i(L)$.

Dans le cas où $\mathcal{L}$ est un cône rationnel nous pouvons donner une autre caractérisation des familles $C_i(\mathcal{L})$ et $CO_i(\mathcal{L})$ (resp. $C_i^f(\mathcal{L})$ et $CO_i^f(\mathcal{L})$) semblable à celle du théorème 3.1. Pour cela définissons :

Définition 3.7

Pour tout langage L et tout ensemble $A \in Rat_2$, nous appellerons

- Crochet de degré i de L par A, le langage

$$C_i(L,A) = \{a_1\ell_1a_2\ldots a_i\ell_i\tilde{b}_i\ell_{i+1}\ldots\tilde{b}_2\ell_{2i-1}\tilde{b}_1 \mid \ell_1\ldots\ell_{2i-1} \in L, \quad (a_1a_2\ldots a_i, b_1b_2\ldots b_i) \in A\}.$$

- Crochet ouvert de degré i de L par A, le langage

$$CO_i(L,A) = \{a_1\ell_1a_2\ldots a_i\ell_ib_1\ell_{i+1}\ldots b_{i-1}\ell_{2i-1}b_i \mid \ell_1\ell_2\ldots\ell_{2i-1} \in L, \quad (a_1a_2\ldots a_i, b_1b_2\ldots b_i) \in A\}.$$

Nous pouvons alors énoncer :

Théorème 3.9

Si $\mathcal{L}$ est un cône rationnel (fidèle) alors $C_i(\mathcal{L})$ ($C_i^f(\mathcal{L})$) et $CO_i(\mathcal{L})$ ($CO_i^f(\mathcal{L})$) sont respectivement égaux à la fermeture par union des familles

$$\mathcal{L}' = \{C_i(L,A) \mid L \in \mathcal{L},\ A \in Rat_2\}$$

$$\mathcal{L}'' = \{CO_i(L,A) \mid L \in \mathcal{L},\ A \in Rat_2\}.$$

Démonstration

Nous ne la ferons que dans le cas de $C_i(\mathcal{L})$ et $C_i^f(\mathcal{L})$, elle est identique pour CO_i et CO_i^f.

Il est clair que pour tout $L \in \mathcal{L}$ le langage $C_i(L)$ appartient à $\mathcal{L}'$. Si nous montrons que $\mathcal{L}_1$ la fermeture par union de $\mathcal{L}'$ est un cône (fidèle), il en résulte que $C_i(\mathcal{L}) \subseteq \mathcal{L}_1$. On démontre de façon identique à ce que nous avons fait pour établir le théorème 3.1 que $\mathcal{L}_1$ est fermé par homomorphisme (non effaçant) et par homomorphisme inverse. Démontrons que $\mathcal{L}_1$ est fermé par intersection avec les langages rationnels :

Soit $K \in Rat(X^*)$ un langage reconnu par un automate fini ayant $Q = \{q_0,\ldots,q_\ell\}$ pour ensemble d'états, q_0 pour état initial et $F \subseteq Q$ pour états finals et δ pour fonction de transition. Soit $J = \{j | q_j \in Q\}$. On supposera que pour tout $j \in J$ on a $\delta(q_j,\varepsilon) = q_j$. Pour tous $m,n \in J$, soient :

$$K_m = \{w \in X^* | \delta(q_0,w) = q_m\}$$

$$K_{m,n} = \{w \in X^* | \delta(q_m,w) = q_n\}$$

$$K'_n = \{w \in X^* | \delta(q_n,w) \in F\}.$$

Tous ces ensembles sont rationnels et $K = \bigcup_{m,n \in J} K_m K_{m,n} K'_n$. Pour toute suite $s = (n_1,n_2,\ldots,n_{4i-2})$ d'indices de J, soient alors

$$A_s = K_{n_1}.K_{n_2,n_3}\ldots K_{n_{2i-2},n_{2i-1}} \times \tilde{K}'_{n_{4i-2}}.\tilde{K}_{n_{4i-4},n_{4i-3}}\ldots\tilde{K}_{n_{2i},n_{2i+1}}$$

et $$K_s = K_{n_1,n_2}.K_{n_3,n_4}\ldots K_{n_{4i-3},n_{4i-2}}.$$

Soit alors L' un langage de $\mathcal{L}'$, $L' = C_i(L,A) \in X^*$, $L \in \mathcal{L}$, $A \in Rat_2$.

On montre alors sans difficulté que

$$L' \cap K = \bigcup_s C_i(L \cap K_s, A \cap A_s)$$

Or $L \cap K_s$ est dans le cône rationnel (fidèle) $\mathcal{L}$ et si

$$A = \{(\varphi h, \psi h) | h \in R\}$$

alors

$$A \cap A_s = \{(\varphi h, \psi h) | h \in R \cap \varphi^{-1}(K_{n_1}.K_{n_2,n_3}\ldots K_{n_{2i-2},n_{2i-1}}) \cap \psi^{-1}(K_{n_{2i},n_{2i+1}}\ldots K_{n_{4i-4},n_{4i-3}}.K'_{n_{4i-2}})\}$$

et donc $A \cap A_s \in Rat_2$.

Chacun des crochets de degré i, $C_i(L \cap K_s, A \cap A_s)$ est donc dans $\mathcal{L}'$ et leur union finie sur s dans $\mathcal{L}_1$. Donc $C_i(\mathcal{L}) \subseteq \mathcal{L}_1$ (resp. $C_i^f(\mathcal{L}) \subseteq \mathcal{L}_1$).

Pour achever la preuve il nous suffit de montrer que tout élément de $\mathcal{L}'$ appartient à $C_i(\mathcal{L})$ (resp. $C_i^f(\mathcal{L})$). Or ce résultat découle d'une preuve analogue à celle du lemme 3.2, que nous ne détaillerons donc pas.

■

Le théorème 3.9 nous permet donc, si l'on remarque que la fermeture par union de $\mathcal{L}'$ est un cône rationnel si $\mathcal{L}$ est un cône rationnel, d'énoncer comme pour le crochet et le crochet ouvert :

Corollaire 3.17

Pour tout langage effaçable L($\mathcal{E}(L) = \mathcal{E}^f(L)$) les langages $C_i(L)$ et $CO_i(L)$ sont effaçables, i.e. $\mathcal{E}^f(C_i(L)) = \mathcal{E}(C_i(L))$ et $\mathcal{E}^f(CO_i(L)) = \mathcal{E}(CO_i(L))$, pour tout $i \geq 1$.

Ce corollaire nous permet donc de construire de nouveaux langages effaçables en général différents de ceux construits à partir du crochet ou du crochet ouvert. Notons cependant que si L est un langage algébrique effaçable alors C(L) est un langage algébrique effaçable, alors qu'en général $C_i(L)$, pour $i \geq 2$, s'il est effaçable n'est plus nécessairement algébrique. En effet si l'on considère le langage algébrique $L_= = \{a^n c b^n | n \in \mathbb{N}\}$ alors $C_2(L)$ est vérifiable [Rod1] mais n'est pas algébrique. En effet on a $C_2(L_=) = \{a_1\ell_1 a_2\ell_2\tilde{b}_2\ell_3\tilde{b}_1 | \ell_1\ell_2\ell_3 \in L_=, a_1a_2 = b_1b_2 \in \{e,d\}^*\}$ et le langage $C_2(L_=) \cap d^*a^*ce^*b^* = \{d^n a^p c e^n b^p | n,p \geq 0\}$ n'est pas algébrique.

Nous allons montrer maintenant que pour tout $i \geq 1$, les opérations C_i, Ch_i, CO_i sont indépendantes de Rat, que le produit de deux langages est indépendant des opérateurs C_i, Ch_i et CO_i (resp. C_i^f, Ch_i^f, CO_i^f), enfin que Ch_i et CO_i sont des opérations syntactiques de degré 0 tandis que C_i n'est syntaxique que pour $i \geq 2$.

Proposition 3.1

Pour tout $i \geq 1$, les opérations C_i, Ch_i et CO_i sont indépendantes de Rat (Rat_0).

Démonstration

Si L est un langage et $\mathcal{L}$ un cône rationnel fermé par union tel que $C_i(L) \in$ Rat $\mathcal{L}$ alors le langage $L' = C_i(L, \{(wc,w) | w \in \{a,b\}^*\}$ qui est rationnellement équivalent à $C_i(L)$ appartient à Rat $\mathcal{L}$ et donc

$L'' = C(L,\{(wc,w)|w \in \{a,b\}^*\} = L' \cap \{a,b\}^* c X_L^* \{a,b\}^*$ appartient à $\underline{\text{Rat}}\ \mathcal{L}$.
D'après le corollaire 3.13 il en résulte que L appartient au cône rationnel $\mathcal{L}$. La preuve est identique dans le cas de CO_i et comme $\underline{\text{Rat}}_0(\mathcal{L}) \subseteq \underline{\text{Rat}}\ \mathcal{L}$, il en résulte que C_i et CO_i sont indépendants de $\underline{\text{Rat}}_0$.

∎

Afin de montrer que le produit de deux langages est indépendant de C_i, Ch_i et CO_i, $i \geq 1$, rappelons la définition et les propriétés d'une opération étudiée dans [Gr3] sous le nom "d'intercalafion of degree i", ainsi que les définitions et propriétés des opérations classiques Init et Fin [Gr1].

Définition 3.8

Pour tout langage $L \subseteq X^*$, soient les langages :

$$\text{Init}(L) = \{w \in X^* | \exists x \in X^*, wx \in L\}$$

$$\text{Fin}(L) = \{w \in X^* | \exists x \in X^*, xw \in L\}.$$

Si L' est un autre langage alors l'opération Shuf_i de L et de L', appelée "Shuffle de degré i" ("intercalation de degré i" dans [Gr3]) est définie par

$$\text{Shuf}_i(L,L') = \{\ell_1 m_1 \ell_2 m_2 \ldots \ell_{i-1} m_{i-1} \ell_i | \ell_1 \ell_2 \ldots \ell_i \in L, m_1 m_2 \ldots m_{i-1} \in L'\}.$$

Le "Shuffle" de L et L' est le langage

$$\text{Shuf}(L,L') = \{\ell_1 m_1 \ldots \ell_n m_n | n \in \mathbb{N}, \ell_1 \ell_2 \ldots \ell_n \in L,\ m_1 m_2 \ldots m_n \in L'\}.$$

On a alors :

Lemme 3.5 [GG2], [Gr3]

Soit $\mathcal{L}$ un cône rationnel (fidèle). Pour tout $L \in \mathcal{L}$ et $R \in Rat$ (Rat_0) les langages Init(L), Fin(L), Shuf(L,R) et $Shuf_i(L,R)$, $i \geq 1$, appartiennent à $\mathcal{L}$.

Si nous associons aux opérations $Shuf_i$, $i \geq 1$ et Shuf les opérateurs suivants :

Définition 3.9

Pour toutes familles de langages $\mathcal{L}$ et $\mathcal{L}'$, soient

$\forall i \geq 1$, $Shuf_i(\mathcal{L}, \mathcal{L}') = \mathcal{C}^U(\{Shuf_i(L,L') | L \in \mathcal{L}, L' \in \mathcal{L}'\})$

et $Shuf(\mathcal{L}, \mathcal{L}') = \mathcal{C}^U(\{Shuf(L,L') | L \in \mathcal{L}, L' \in \mathcal{L}'\})$.

Nous pouvons alors reformuler les résultats de [Gr3] sous la forme :

Proposition 3.2 [Gr3]

Si $\mathcal{L}$ et $\mathcal{L}'$ sont deux cônes rationnels, alors :

1) Pour tout $i \geq 1$, $Shuf_i(\mathcal{L}, \mathcal{L}')$ est égal à la fermeture par union de la famille $\mathcal{L}_i = \{Shuf_i(L,L') \; L \in \mathcal{L}, L' \in \mathcal{L}'\}$.

2) $Shuf_i$ est une opération régulière, syntaxique de degré i supérieure à $Shuf_{i-1}$ si $i > 2$, supérieure à l'opérateur $\odot$ si $i = 2$. Si $i = 1$ alors $Shuf_1(\mathcal{L}, \mathcal{L}') = \mathcal{L} \odot \mathcal{L}'$.

Nous pouvons alors énoncer :

Théorème 3.10

Le produit de deux langages est indépendant des opérateurs C_i, Ch_i

et CO_i (resp. C_i^f, Ch_i^f, CO_i^f) pour tout $i \geq 1$.

Démonstration

Puisque $C_i^f(\mathcal{L}) \subseteq C_i(\mathcal{L})$, $Ch_i^f(\mathcal{L}) \subseteq Ch_i(\mathcal{L})$ et $CO_i^f(\mathcal{L}) \subseteq CO_i(\mathcal{L})$, le résultat pour les opérateurs fidèles résulte directemen de celui pour les opérateurs C_i, Ch_i et CO_i. De même comme $Ch_i(\mathcal{L}) \subseteq C_i(\mathcal{L})$, il nous suffit d'établir le résultat pour C_i et CO_i. Traitons tout d'abord le cas de l'opérateur C_i.

Soient donc L_1 et L_2 deux langages sur des alphabets disjoints X_1 et X_2 et $\mathcal{L}$ un cône rationnel fermé par union.

Si $L_1L_2 \in C_i(\mathcal{L})$, d'après le théorème 3.9, il existe un entier n, n ensembles $A_j \in Rat_2$ et n langages $M_j \in \mathcal{L}$ tels que :

$$L_1L_2 = \bigcup_{j=1}^{n} C_i(M_j,A_j).$$

Nous supposerons que A_j et M_j sont non vides pour tout j. Pour tout $w_1 \in L_1$ et $w_2 \in L_2$ il existe donc j, $1 \leq j \leq n$, $(a_1\ldots a_i, b_1\ldots b_i) \in A_j$ et $\ell = \ell_1\ldots\ell_{2i-1} \in M_j$ tels que

$$w_1w_2 = a_1\ell_1a_2\ell_2 \ldots a_i\ell_i\tilde{b}_i \ldots \ell_{2i-1}\tilde{b}_1.$$

Deux cas se présentent :

1er cas

$\forall w_1 \in L_1$, $\exists w_2 \in L_2$, $\exists j$, $1 \leq j \leq n$, $(a_1\ldots a_i, b_1\ldots b_i) \in A_j$ et $\ell = \ell_1\ldots\ell_{2i-1} \in M_j$ tels que :

$$w_1w_2 = a_1\ell_1a_2\ell_2\ldots a_i\ell_i\tilde{b}_i\ldots\ell_{2i-1}\tilde{b}_1$$

et w_1 est facteur gauche propre de $a_1\ell_1a_2\ldots\ell_{i-1}a_i$. Soit alors

$$L = \{wf \in X_1^* \mid w = a_1\ell_1a_2\ldots\ell_{p-1}a_p,\ 1 \le p \le i,\ \exists f' \in X_2^*X_2,$$

$$a_{p+1},\ldots,a_i,b_1,\ldots,b_i \in X_2^*,\ 1 \le j \le n : ff' = \ell_p,$$

$$(a_1\ldots a_i,b_1\ldots b_i) \in A_j,\ \ell_1\ldots\ell_p \in \mathrm{Init}(M_j)\}.$$

Montrons que $L_1 = L$. Les conditions de ce premier cas impliquent $L_1 \subseteq L$. Réciproquement soit $wf \in L$. Alors il existe j, $1 \le j \le n$, $p < i$, $\ell = \ell_1\ldots\ell_p\ldots\ell_{2i-1} \in M_j$, $(a_1\ldots a_p\ldots a_i,b_1\ldots b_i) \in A_j$, $f' \in X_2X_2^*$ tels que

$$w = a_1\ell_1 \ldots \ell_{p-1}a_p,\ ff' = \ell_p$$

Alors

$$wff'\ a_{p+1}\ \ell_{p+1}\ldots a_i\ell_i\tilde{b}_i\ldots\ell_{2i-1}\tilde{b}_1 \in C_i(M_j,A_j) \subseteq L_1L_2.$$

La disjonction des alphabets X_1 et X_2 implique donc que $wf \in L_1$ et donc $L \subseteq L_1$.

Soit alors $\pi_1(A_j) = \{a_1\ldots a_i \mid \exists b_1,\ldots,b_i : (a_1\ldots a_i,b_1\ldots b_i) \in A_j\}$.

Soit $P_1(A_j) = \{w \in X_1^* \mid \exists f \in X_2^* : wf \in \pi_1(A_j)\}$.

$\pi_1(A_j)$ est un langage rationnel, donc $P_1(A_j)$ qui est obtenu à partir de $\pi_1(A_j)$ par une transduction rationnelle facile à construire, est aussi rationnel. De même $M_j' = \{w \in X_1^* \mid \exists f \in X_2X_2^* : wf \in M_j\}$ appartient à $\mathcal{L}$ puisque $M_j \in \mathcal{L}$. On a alors de façon évidente :

$$L_1 = L = \bigcup_{j=1}^{n} \mathrm{Shuf}_{i-1}(P_1(A_j),M_j')$$

et donc d'après le lemme 3.5, $L_1 \in \mathcal{L}$ puisque $\mathcal{L}$ est un cône rationnel fermé par union.

2ème cas

$\exists w_1 \in L_1$ tel que pour tout $w_2 \in L_2$, $\forall j$, $1 \leq j \leq n$,

$(a_1 \ldots a_i, b_1 \ldots b_i) \in A_j$, $\ell = \ell_1 \ldots \ell_{2i-1} \in M_j$ tels que

$w_1 w_2 = a_1 \ell_1 \ldots a_i \ell_i \tilde{b}_i \ldots \ell_{2i-1} \tilde{b}_1$ implique que $a_1 \ell_1 \ldots \ell_{i-1} a_i$

est facteur gauche de w_1.

Soit alors

$L(w_1) = \{fw \in X_2^* \mid \exists j,\ 1 \leq j \leq n,\ \exists p,\ 1 \leq p \leq i,\ \exists f',$

$a_1, \ldots, a_i, b_{p+1}, b_{p+2}, \ldots, b_i \in X_1^* : w = \tilde{b}_p \ell_{2i-p+1} \tilde{b}_{p-1} \ldots \ell_{2i-1} \tilde{b}_1,$

$f'f = \ell_{2i-p}$, $(a_1 a_2 \ldots a_i, b_1 b_2 \ldots b_i) \in A_j$,

$\ell_{2i-p} \ell_{2i-p+1} \ldots \ell_{2i-1} \in \mathrm{Fin}(M_j), a_1 \ell_1 \ldots \ell_{2i-1} a_i \tilde{b}_i \ldots \tilde{b}_{p+1} f = w_1\}$.

Montrons que $L_2 = L(w_1)$. Les conditions caractérisant ce deuxième cas impliquent $L_2 \subseteq L(w_1)$. Réciproquement soit $fw \in L(w_1)$. Alors il existe j, $1 \leq j \leq n$, $(a_1 \ldots a_i, b_1 \ldots b_i) \in A_j$, $\ell_1 \ldots \ell_{2i-1} \in M_j$ et p, $1 \leq p \leq i$, $f' \in X_1^*$ tels que $f'f = \ell_p$, $w = \tilde{b}_p \ldots \ell_{2i-1} \tilde{b}_1$ et donc

$$\ell = a_1 \ell_1 \ldots a_i \ell_i \tilde{b}_i \ldots \tilde{b}_{p+1} f' fw = a_1 \ell_1 \ldots a_i \ell_i \tilde{b}_i \ldots \ell_{2i-1} \tilde{b}_1$$

est un mot de $C_i(M_j, A_j)$, donc $\ell \in L_1 L_2$. La disjonction des alphabets X_1 et X_2 implique donc que $fw \in L_2$, soit $L(w_1) \subseteq L_2$.

Quel que soit le facteur droit w_1' de w_1 et quel que soit p, $1 \leq p \leq i$, le nombre de factorisation de w_1' en $a_1\ell_1\ldots\tilde{b}_{p+1}\ell_{2i-p}$ est fini. Pour une factorisation σ_p donnée, soit :

$$P_j(\sigma_p,w_1') = \{w \in X_2^* | (a_1\ldots a_i, wb_{p+1}\ldots b_i) \in A_j\}$$

$$M_j(\sigma_p,w_1') = \{w \in X_2^* | \ell_1\ldots\ell_{2i-p}\, w \in M_j\}.$$

$P_j(\sigma,w_1')$ est rationnel puisque obtenu à partir de $\pi_2(A_j)$ par la transduction rationnelle qui rejette les mots ne finissant pas par $b_{p+1}\ldots b_i$ et efface ce suffixe aux autres. De même $M_j(\sigma,w_1') \in \mathcal{L}$ puisqu'obtenu à partir de M_j par la transduction qui rejette les mots ne commençant pas par $\ell_1\ldots\ell_{2i-p}$ et efface ce préfixe aux autres. Or nous avons clairement :

$$L_2 = L(w_1) = \bigcup_{w_1'} \bigcup_{\sigma_p} \mathrm{Shuf}_p(\tilde{P}_j(\sigma_p,w_1'),M_j(\sigma_p,w_1'))$$

Donc d'après le lemme 3.5, $L_2 \in \mathcal{L}$ puisque $\mathcal{L}$ est fermé par union, et $\tilde{P}_j(\sigma_p,w_1')$ est rationnel.

Pour montrer que $L_1L_2 \in CO_i(\mathcal{L})$ implique L_1 ou $L_2 \in \mathcal{L}$, la preuve est identique dans le premier cas, tandis que dans le second

$$L_2 = \bigcup_{w_1'} \bigcup_{\sigma_p} (P_j'(\sigma_p',w_1'),M_j(\sigma_p',w_1')), \text{ avec}$$

$$P_j'(\sigma_p',w_1') = \{w \in X_2^* | (a_1\ldots a_i, b_1\ldots b_p w) \in A_j\}$$

$$M_j(\sigma_p',w_1') = \{w \in X_2^* | \ell_1\ldots\ell_{2i-p} w \in M_j\}$$

pour toute factorisation σ_p' de w_1' en $a_1\ell_1\ldots a_i\ell_i b_1\ell_{i+1}\ldots b_p\ell_{2i-p}$.

■

Lemme 3.6

1) L'opération chevron est indépendante de $CO_i(CO_i^f)$ pour tout $i \geq 1$.

2) Pour tout langage L et tout cône rationnel (fidèle) $\mathcal{L}$ la condition $Ch(L) \in C_i(\mathcal{L})$ (resp. $C_i^f(\mathcal{L})$) implique que $Ch(L) \in \mathcal{L} \cup C(Rat) = \mathcal{L} \cup Lin$.

Démonstration

Si $i = 1$, le résultat découle du théorème 3.3. Supposons $i \geq 2$ et soit L un langage tel que $X_L \cap \{a,b\} = \emptyset$ et $\mathcal{L}$ un cône rationnel tels que $Ch(L) = \{a^n \ell b^n \mid n \in \mathbb{N}, \ell \in L\}$ appartient à $CO_i(\mathcal{L})$. Si $L = \{\varepsilon\}$ alors trivialement $L \in \mathcal{L}$. Supposons donc que $L \neq \{\varepsilon\}$. D'après le théorème 3.9 il existe donc un entier $p \in \mathbb{N}$, p langages $M_j \in \mathcal{L}$ et p ensembles $A_j \in Rat_2$ tels que $Ch(L) = \bigcup_{j=1}^{p} CO_i(M_j,A_j)$. Donc pour tout $n \in \mathbb{N}$ et tout $\ell \in L$, il existe j, $1 \leq j \leq p$, un mot $\ell_1\ell_2\ldots\ell_{2i-1} \in M_j$ et $(a_1a_2\ldots a_i, b_1b_2\ldots b_i) \in A_j$ tels que $a^n \ell b^n = a_1\ell_1\ldots a_i\ell_i b_1\ell_{i+1}\ldots\ell_{2i-1}b_i$. Nous pouvons supposer que A_j et M_j sont non vides pour tout j, $1 \leq j \leq p$. Montrons que si $M_j \neq \{\varepsilon\}$ alors $A_j = \{(\varepsilon,\varepsilon)\}$. En effet sinon soit $m \neq \varepsilon$ un mot de M_j et pour tout $(a_1,b_1) \in A_j$ on a $(a_1,b_1) \neq (\varepsilon,\varepsilon)$. Supposons que $m \in a^+X_L^*$. Si a_1 et b_1 sont différents de ε alors, par exemple, $a_1 \in a^+$ et $b_1 \in X_L^*b^+$. Il en résulte que le mot $a_1b_1m \in a^+X_L^*b^+a^+X_L^*$ appartiendrait à Ch(L) ce qui est contradictoire. Les autres cas conduisent, de même, à une contradiction. Si $a_1 \neq \varepsilon$ et $b_1 = \varepsilon$, alors $a_1m \in a^+X_L$ appartiendrait à Ch(L), ce qui est impossible. Si l'on suppose que $m \in a^*X_L^+b^*$ ou $m \in a^*X_L^*b^+$, on aboutit de même à une contradiction.

Il résulte de ceci que l'on a donc

$$Ch(L) = \bigcup_{j,M_j \neq \{\varepsilon\}} M_j \cup \left(\bigcup_{j,M_j = \{\varepsilon\}} \{a_j b_j | (a_j,b_j) \in A_j\} \right).$$

Ainsi, si pour tout j, $1 \leq j \leq p$, on a $M_j \neq \varepsilon$, le langage $Ch(L) = \bigcup_{j=1}^{p} M_j$ appartient à $\mathcal{L}$ et donc L appartient à $\mathcal{L}$, sinon Ch(L) appartient à $\mathcal{L} \cup CO(Rat)$ qui est inclus dans $CO(\mathcal{L})$. D'après le théorème 3.3 on sait que Ch est indépendant de CO, il en résulte donc que $L \in \mathcal{L}$, ce qui prouve le point 1) du lemme.

Dans le cas de l'opérateur C_i le raisonnement précédent reste vrai mais nous ne pouvons que conclure que Ch(L) appartient à $\mathcal{L} \cup C(Rat)$. D'où le point 2) du lemme.

■

Ceci nous permet alors d'énoncer :

Théorème 3.11

Pour tout $i \geq 1$, les opérations Ch_i, CO_i sont syntactiques de degré 0. C_i est de degré 0 si $i \geq 2$.

Démonstration

Nous avons vu que pour tout $i \geq 1$, Ch_i, CO_i et C_i sont indépendantes de Rat (Proposition 3.1). D'autre part si $CO_i(L) \in CO_i(\mathcal{L})$ pour quelque langage L et quelque cône rationnel L formé par union, alors Ch(L) appartient à $CO_i(\mathcal{L})$ et le résultat découle du lemme précédent.

Comme $Ch_i(\mathcal{L}) \subseteq CO_i(\mathcal{L})$ il en résulte que Ch_i est également de degré 0.

Supposons maintenant que L est un langage et $\mathcal{L}$ un cône rationnel

tels que $C_i(L) \in C_i(\mathcal{L})$. Comme

$$\{yx\tilde{x} \mid y \in L, x \in \{a,b\}^*\} \subseteq C_i(L)$$

il en résulte que $L.S_2 = C_i(L) \cap X_L^* \{a,b\}^*$ appartient à $C_i(\mathcal{L})$. D'après le théorème 3.10 il en résulte que L ou S_2 est dans $\mathcal{L}$. D'autre part, comme la condition $C_i(L) \in C_i(\mathcal{L})$ implique $Ch(L) \in C_i(\mathcal{L})$, il résulte du lemme 3.6 que $L \in \mathcal{L} \cup Lin$. Si donc $Lin \not\subseteq \mathcal{L}$ alors $S_2 \notin \mathcal{L}$ et donc $L \in \mathcal{L}$ et si $Lin \subseteq \mathcal{L}$ alors on a également $L \in \mathcal{L}$. Donc C_i est de degré 0 si $i \geq 2$.

■

Des propositions 3.1, théorème 3.10 et théorème 3.11 il résulte alors, d'après les résultats du chapitre 2 :

Corollaire 3.18

Soit $\mathcal{L}$ un cône rationnel (fidèle) fermé par union et non fermé par chevron de degré $i \geq 1$ (resp. Ch_i^f, CO_i, CO_i^f et C_i et C_i^f pour $i \geq 2$) alors

1) Le plus petit cône rationnel (fidèle) contenant $\mathcal{L}$ et fermé par Ch_i (resp. Ch_i^f, CO_i, CO_i^f et CO_i et C_i^f si $i \geq 2$) est non principal.

2) Le plus petit cône rationnel (fidèle) contenant $\mathcal{L}$ et fermé par produit et par Ch_i (resp. Ch_i^f, CO_i, CO_i^f et C_i et C_i^f, $i \geq 2$) est non principal.

3) La plus petite FAL (fidèle) contenant $\mathcal{L}$ et fermée par Ch_i (resp. Ch_i^f, CO_i, CO_i^f et C_i et C_i^f pour $i \geq 2$) est non principale.

4) Pour $t = Ch_i$ (resp. Ch_i^f, CO_i, CO_i^f et C_i et C_i^f pour $i \geq 2$), on a :

$$\mathcal{E}_t^U(\mathcal{L}) \subsetneq \underline{\text{Rat}}\ \mathcal{E}_t^U(\mathcal{L}) \subsetneq \mathcal{F}_t(\mathcal{L})$$

$$(\mathcal{E}_t^{fU}(\mathcal{L}) \subsetneq \underline{\text{Rat}}\ \mathcal{E}_t^{fU}(\mathcal{L}) \subsetneq \mathcal{F}_t^f(\mathcal{L}))$$

$$\mathcal{E}_t^U(\mathcal{L}) \subsetneq \underline{\text{Fin}}\ \mathcal{E}_t^U(\mathcal{L}) \subsetneq \mathcal{E}_t^{U\bullet}(\mathcal{L})$$

$$(\mathcal{E}_t^{fU}(\mathcal{L}) \subsetneq \underline{\text{Fin}}\ \mathcal{E}_t^{fU}(\mathcal{L}) \subsetneq \mathcal{E}_t^{fU\bullet}(\mathcal{L})).$$

Remarque

On sait [BCN] que l'opérateur crochet est idempotent et l'opération crochet est un cas particulier d'insertion d'un langage dans S_2 le générateur de Lin, la famille des langages linéaires. De même l'opération crochet de degré i est un cas particulier de l'opération Shuf_{2i} ("intercalation of degree 2i") de L et S_2. Or cette opération est de degré 2i [Gr3] et n'est pas de degré 0, tandis que l'opération C_i est de degré 0. Ceci, compte tenu de l'étude d'un certain nombre d'opérations unaires nous conduit à poser la conjecture :

Conjecture 3.1

Toute opération unaire de degré p est de degré 0.

Le lemme 3.6 nous permet également de tirer

Corollaire 3.19

Si $\mathcal{L}$ est un cône rationnel (fidèle) fermé par union et L un langage tel que $C(L) \in \mathcal{E}_{CO_i}^U(\mathcal{L})$, $i \geq 1$, alors $C(L) \in \mathcal{L}$.

Démonstration

Si $C(L) \in \mathcal{E}_{CO_i}^U(\mathcal{L})$, alors il existe $p \geq 0$ tel que

$C(L) \in CO_i^P(\mathcal{L})$ et, comme l'opérateur crochet est idempotent il en résulte que $C(C(...C(L))) \in CO_i^P(\mathcal{L})$ et donc $Ch(Ch(.Ch(C(L))...)) \in CO_i^P(\mathcal{L})$ et l'application p fois répétée du lemme 3.6 implique que $C(L) \in \mathcal{L}$. ∎

Nous terminerons cette section par deux lemmes qui vont nous permettrent de caractériser les plus petits cônes rationnels (fidèles) clos par crochet et crochet ouvert de degré i :

Lemme 3.7

Soit $\mathcal{L}$ un cône rationnel (fidèle) alors pour tout $i \geq 1$:

1) $Shuf_i(C(Rat), \mathcal{L}) \subseteq C_i(\mathcal{L}) \subseteq Shuf_{2i}(C(Rat), \mathcal{L})$

2) $Shuf_i(CO(Rat), \mathcal{L}) \subseteq CO_i(\mathcal{L}) \subseteq Shuf_{2i}(CO(Rat), \mathcal{L})$.

Démonstration

Les preuves pour C_i et CO_i sont semblables.

a) D'après le corollaire 3.16, $C_1(Rat) = \mathcal{C}(S_2)$ où $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$. Donc il nous suffit de vérifier que $Shuf_i(S_2,L) \in C_i(\mathcal{L})$, $\forall L \in \mathcal{L}$. On peut supposer $X_L \cap X_{S_2} = \emptyset$. Or :

$$Shuf_i(S_2,L) = \{a_1\ell_1...a_{i-1}\ell_{i-1}a_i | a_1a_2...a_i \in S_2, \ell_1...\ell_{i-1} \in L\}$$

et
$$C_i(L) = \{a_1'\ell_1'...a_i'\ell_i'\tilde{b}_i...\ell_{2i-1}'\tilde{b}_1 | a_1'...a_i'\tilde{b}_i...\tilde{b}_1 \in S_2, \ell_1...\ell_{2i-1} \in L\}.$$

Donc

$Shuf_i(S_2,L) = C_i(L) \cap (X_{S_2}^* X_L^*)^{i-1} X_{S_2}^*$. D'où le résultat.

b) De la même façon

$$C_i(L) = \mathrm{Shuf}_{2i}(S_2, L) \cap (\{a,b\}^* X_L^*)^{i-1} \{a,b\}^* c X_L^* (\{a,b\}^* X_L)^{i-1} \{a,b\}^*.$$

■

Il en résulte alors :

Théorème 3.12

Soit $\mathcal{L}$ un cône rationnel. Alors une condition nécessaire et suffisante pour que, $\forall i > 1$ on ait :

a) $C_1(\mathcal{L}) = C_i(\mathcal{L})$ est que $C_1(\mathcal{L}) = \mathrm{Shuf}^U_{2p}(C_1(\mathrm{Rat}), \mathcal{L})$ pour tout $p \geq \mathbb{N}^+$.

b) $CO_1(\mathcal{L}) = CO_i(\mathcal{L})$ est que $CO_1(\mathcal{L}) = \mathrm{Shuf}^U_{2p}(CO_1(\mathrm{Rat}), \mathcal{L})$ pour tout $p \in \mathbb{N}^+$.

Démonstration

Celle de b) est identique à celle de a). Si pour tout $p \geq 1$ on a $C_1(\mathcal{L}) = \mathrm{Shuf}^U_{2p}(C_1(\mathrm{Rat}), \mathcal{L})$ alors d'après le lemme 3.7, $C_i(\mathcal{L}) \subseteq C_1(\mathcal{L})$ et l'inclusion réciproque est triviale.

Réciproquement si $C_1(\mathcal{L}) = C_i(\mathcal{L})$, $\forall i \geq 1$, alors d'après le lemme 3.7 :

$$C_1(\mathcal{L}) = C_i(\mathcal{L}) \subseteq \mathrm{Shuf}^U_{2i}(C_1(\mathrm{Rat}), \mathcal{L}) \subseteq C_{2i}(\mathcal{L}) = C_1(\mathcal{L})$$

D'où le résultat.

■

Comme Shuf($\mathcal{L}$,Rat) $\subseteq$ $\mathcal{L}$ il en résulte immédiatement en tenant compte du corollaire 3.17 :

Corollaire 3.20

Pour tout $i \geq 1$:

1) $C(Rat) = C^f(Rat) = C_i(Rat) = C_i^f(Rat) = \mathcal{E}(S_2) = \mathcal{E}^f(S_2)$

2) $CO(Rat) = CO^f(Rat) = CO_i(Rat) = CO_i^f(Rat) = \mathcal{E}(L_1) = \mathcal{E}^f(L_1)$

avec $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et $L_1 = \{wcw | w \in \{a,b\}^*\}$.

C) LES OPERATIONS CROCHET ET CROCHET OUVERT DE DEGRE INFINI

Le crochet (resp. crochet ouvert) de degré infini d'un langage L n'est autre que $\bigcup_{i \geq 1} C_i(L)$ (resp. $\bigcup_{i \geq 1} CO_i(L)$). Cette opération notée C_∞ (resp. CO_∞) peut donc être considérée comme la limite de la suite des opérations $\{C_i\}_{i \in \mathbb{N}}$ (resp. $\{CO_i\}_{i \in \mathbb{N}}$). Cette démarche a été déjà entreprise par S. Greibach dans [Gr3] où l'on établit que la suite des opérations syntaxiques binaires $\{Shuf_i\}_{i \in \mathbb{N}}$ a pour limite l'opération non syntaxique Shuffle.

Nous montrons tout d'abord que ce passage à la limite préserve l'indépendance par rapport à l'opérateur Rat. Utilisant alors les propriétés de l'opérateur Shuf et de l'opérateur $\mathcal{H}$ (resp. $\mathcal{H}^f$) de fermeture par homomorphisme (resp. fidèle) nous montrons que l'application trois fois répétée des opérateurs (fidèles) associés à CO_∞ et C_∞ conduit à la fermeture, pour ces opérations, de tout cône rationnel (fidèle). Nous établissons ainsi que $\mathcal{L}_{BNP}$ [BNP] et $\mathcal{L}_{DUP}$ [BG] sont les plus petits cônes rationnels fidèles fermés respectivement par C_∞ et CO_∞. De même RE, la famille des langages récursivement énumérables, est le plus petit cône rationnel fermé par C_∞ et la plus petite FAL fermée par CO_∞. Le problème reste ouvert de savoir si $\mathcal{C}_{CO_\infty}(Rat)$, le plus petit cône rationnel fermé par intersection et CO_∞, est également la famille RE, de même que reste à démontrer l'inclusion stricte $\mathcal{L}_{DUP} \not\subseteq \mathcal{L}_{BNP}$ ([BN,BG]).

Définition 3.10

Soit $L \subseteq X_L^*$ un langage et $X = \{a,b\}$ un alphabet disjoint de X_L et c un nouveau symbole. Nous appellerons

- Crochet de degré infini de L, le langage

$$C_\infty(L) = \{x_1y_1x_2y_2\ldots x_ny_n \mid n \in \mathbb{N}, x_1x_2\ldots x_n \in L, y_1y_2\ldots y_n \in \{wc\hat{w} \mid w \in X^*\}\}.$$

- Crochet ouvert de degré infini de L, le langage

$$CO_\infty(L) = \{x_1y_1x_2y_2\ldots x_ny_n \mid n \in \mathbb{N}, x_1x_2\ldots x_n \in L, \quad y_1y_2\ldots y_n \in \{wcw \mid w \in X^*\}\}$$

comme les opérations crochet et crochet ouvert de degré i, $C_\infty(L)$ et $CO_\infty(L)$ sont définis à un isomorphisme alphabétique près. Il est clair que l'on a

$$C_\infty(L) = \mathrm{Shuf}(L, S_2),\ S_2 = \{wc\tilde{w} \mid w \in X^*\}$$

et $$CO(L) = \mathrm{Shuf}(L, L_1),\ L_1 = \{wcw \mid w \in X^*\}.$$

Notations

Pour toute famille de langages $\mathcal{L}$ nous désignerons par $C_\infty(\mathcal{L})$ (resp. $C^f_\infty(\mathcal{L})$, $CO_\infty(\mathcal{L})$, $CO^f_\infty(\mathcal{L})$) le plus petit cône rationnel (fidèle) fermé par union contenant la famille $\mathcal{L}' = \{C_\infty(L) \mid L \in \mathcal{L}\}$ ($\mathcal{L}'' = \{CO_\infty(L) \mid L \in \mathcal{L}\}$.

Rappelons enfin les notations introduites dans les sections précédentes :

- Pour toutes familles de langages $\mathcal{L}_1$ et $\mathcal{L}_2$

$$\mathcal{L}_1 \wedge \mathcal{L}_2 = \{L_1 \cap L_2 \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\}$$

$\mathcal{H}(\mathcal{L}_1)$ (resp. $\mathcal{H}^f(\mathcal{L}_1)$) désigne la fermeture par homomorphisme (resp. non effaçant) de $\mathcal{L}$.

$$\mathrm{Shuf}(\mathcal{L}_1, \mathcal{L}_2) = \mathcal{E}^U(\{\mathrm{Shuf}(L_1, L_2) \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\}).$$

Notons enfin que pour tous langages L_1, L_2, L_3, nous avons

$$Shuf(L_1, Shuf(L_2,L_3)) = Shuf(Shuf(L_1,L_2),L_3)$$

que nous noterons donc $Shuf(L_1,L_2,L_3)$.

Nous énoncerons tout d'abord

Proposition 3.3

C_∞ et CO_∞ sont des opérations indépendantes de Rat (Rat_0).

Démonstration

Elle est identique pour C_∞ et CO_∞. Si L est un langage et $\mathcal{L}$ un cône rationnel fermé par union tels que $C_\infty(L) \in$ Rat $\mathcal{L}$ alors $C(L) = C_\infty(L) \cap \{a,b\}^* \, cX_L^* \, \{a,b\}^*$ appartient à Rat $\mathcal{L}$ et donc $L \in \mathcal{L}$ puisque l'opération crochet C est indépendante de Rat.

■

Nous allons maintenant montrer que les opérations C_∞ (resp. C_∞^f, CO_∞ et CO_∞^f) ne sont pas de degré $p \in \mathbb{N}$ et que les opérateurs associés C_∞ (resp. C_∞^f, CO_∞ et CO_∞^f) ne sont pas idempotents.

Rappelons tout d'abord un certain nombre de résultats :

Proposition 3.4 [Gi2]

Soient L_1 et L_2 deux langages non vides sur des alphabets disjoints. Alors

1) $\mathcal{H}^f(\mathcal{E}^f(L_1) \wedge \mathcal{E}^f(L_2)) = \mathcal{E}^f(Shuf(L_1,L_2))$

2) $\mathcal{H}(\mathcal{E}^f(L_1) \wedge \mathcal{E}^f(L_2)) = \mathcal{E}(\text{Shuf}(L_1,L_2))$

Définition 3.11 [BNP]

$\mathcal{L}_{BNP}$ désignera le plus petit cône rationnel fidèle fermé par union et intersection contenant les langages linéaires. $\mathcal{L}_{BNP}$ est égal à $\mathcal{H}^f(\text{Lin} \wedge \text{Lin} \wedge \text{Lin})$ d'après [BNP].

Définition 3.12 [B.G.W] [B.G]

$\mathcal{L}_{DUP}$ désignera le plus petit cône rationnel fidèle fermé par union et intersection contenant le langage $L_1 = \{wcw | w \in \{a,b\}^*\}$. D'après [BGW] tout langage de L_{DUP} est image par un homomorphisme non effaçant de trois langages appartenant à $\mathcal{E}(L_1)$.

On a alors :

Proposition 3.5 [BN] [BGW]

Pour tout cône rationnel fidèle $\mathcal{L}$, les familles $\mathcal{H}^f(\mathcal{L} \wedge \mathcal{L}_{BNP})$ et $\mathcal{H}^f(\mathcal{L} \wedge \mathcal{L}_{DUP})$ sont fermées par intersection.

Nous pouvons alors énoncer :

Théorème 3.13

Pour tout cône rationnel (fidèle) $\mathcal{L}$:

1) $C_\infty^f(\mathcal{L}) = \mathcal{H}^f(\text{Lin} \wedge \mathcal{L})$ (resp. $CO_\infty^f(\mathcal{L}) = \mathcal{H}^f(\mathcal{L} \wedge \mathcal{E}(L_1))$

2) $C_\infty^{f2}(\mathcal{L}) = \mathcal{H}^f(\text{Lin} \wedge \text{Lin} \wedge \mathcal{L})$
(resp. $CO_\infty^{f2}(\mathcal{L}) = \mathcal{H}^f(\mathcal{L} \wedge \mathcal{E}(L_1) \wedge \mathcal{E}(L_1))$

3) Pour tout $p \geq 3$

$$C_\infty^{fp}(\mathcal{L}) = \mathcal{HC}^f(\mathcal{L}_{BNP} \wedge \mathcal{L})$$

$$(\text{resp. } CO_\infty^{fp}(\mathcal{L}) = \mathcal{HC}^f(\mathcal{L}_{DUP} \wedge \mathcal{L})).$$

Démonstration

Elle est identique pour C_∞^f et CO_∞^f.

Par définition

$$C_\infty^f(\mathcal{L}) = \mathcal{E}^{fU}(\{C_\infty(L) | L \in \mathcal{L}\}) = \mathcal{E}^{fU}(\{Shuf(S_2,L) | L \in \mathcal{L}\})$$

d'où le résultat d'après la proposition 3.4.

L'associativité de l'opérateur $\mathcal{HC}^f$ implique alors que pour tout $p \geq 2$, $C^{fp}(\mathcal{L}) = \mathcal{HC}^f(Lin \wedge \ldots \wedge Lin \wedge \mathcal{L})$ et d'après [BNP] si $p \geq 3$ on a bien $C_\infty^{fp}(\mathcal{L}) = \mathcal{HC}^f(\mathcal{L}_{BNP} \wedge \mathcal{L})$.

■

En particulier si $\mathcal{L}$ est un cône rationnel (fidèle) principal, nous avons :

Corollaire 3.21

Soit $\mathcal{L}$ un cône rationnel (fidèle) principal de générateur L, alors $C_\infty^f(\mathcal{L})$, $C_\infty^{f2}(\mathcal{L})$ et $\mathcal{E}_{C_\infty^f}(\mathcal{L})$ sont des cônes rationnels fidèles principaux de générateurs respectifs $Shuf(L,S_2)$, $Shuf(L,S_2,S_2')$ et $Shuf(L,S_2,S_2',S_2'')$ où S_2' et S_2'' sont des recopies de S_2.

De la même façon $CO_\infty^f(\mathcal{L})$, $CO_\infty^{f2}(\mathcal{L})$ et $\mathcal{E}_{CO_\infty^f}(\mathcal{L})$ sont principaux de générateurs respectifs $Shuf(L,L_1)$, $Shuf(L,L_1,L_2')$ et $Shuf(L,L_1,L_1',L_1'')$ où L_1' et L_1'' sont des recopies de L_1.

Dans le cas où $\mathcal{L}$ = Rat, nous obtenons :

Corollaire 3.22

1) $\mathcal{L}_{BNP}$ est le plus petit cône rationnel fidèle fermé par C_∞.
$\mathcal{L}_{BNP}$ est principal de générateur $Shuf(S_2,S_2',S_2'')$.

2) $\mathcal{L}_{DUP}$ est le plus petit cône rationnel fidèle fermé par CO_∞.
$\mathcal{L}_{DUP}$ est principal de générateur $Shuf(L_1,L_1',L_1'')$.

Corollaire 3.23

1) Pour tout cône rationnel $\mathcal{L}$ inclus dans $\mathcal{L}_{BNP}$, on a

$$C_\infty^{fp}(\mathcal{L}) = \mathcal{L}_{BNP}, \text{ pour tout } p \geq 3.$$

2) Pour tout cône rationnel $\mathcal{L}$ inclus dans $\mathcal{L}_{DUP}$, on a

$$CO_\infty^{fp}(\mathcal{L}) = \mathcal{L}_{DUP}, \text{ pour tout } p \geq 3.$$

Dans le cas des opérateurs C_∞ et CO_∞ nous pouvons seulement énoncer

Corollaire 3.24 [GHa]

Le plus petit cône rationnel fermé par C_∞ est la famille RE des langages récursivement énumérables.

De même comme $L = \{a^n b^n \mid n \in \mathbb{N}\}$ appartient à $CO_\infty(\mathcal{L})$ pour tout cône rationnel $\mathcal{L}$ on a d'après [H.H.].

Corollaire 3.25

La plus petite FAL fermée par CO_∞ est la famille RE des langages récursivement énumérables.

En fait on sait [BGW] que $\mathcal{L}_{DUP}$ est une FAL fidèle, il en résult donc que $\mathcal{L}_{DUP}$ est la plus petite FAL fidèle fermée par CO_∞. Le problème reste ouvert de savoir si RE est également le plus petit cône rationnel fermé par CO_∞. De même l'égalité $\mathcal{L}_{DUP} = \mathcal{L}_{BNP}$ n'a pas été démontrée. On sait que $\mathcal{L}_{DUP} \subseteq \mathcal{L}_{BNP}$ et que l'inclusion réciproque a lieu si et seulement si $S_2 \in \mathcal{L}_{DUP}$ [BGW]. Or, d'autre part, Book et Nivat conjecturent dans [BN] que $\mathcal{L}_{BNP}$ n'est pas fermé par étoile propre : il semblerait donc que l'on ait l'inclusion stricte $\mathcal{L}_{DUP} \neq \mathcal{L}_{BNP}$.

CHAPITRE 4

RELATIONS RATIONNELLES N-AIRES

A) INTRODUCTION

B) DEFINITIONS ET PROPRIETES ELEMENTAIRES

C) LEMME D'ITERATION

D) RELATIONS RATIONNELLES PARTICULIERES

E) RELATIONS RATIONNELLES ET APPLICATIONS

A) INTRODUCTION

Afin d'étudier, dans le chapitre 5, deux familles d'opérations syntaxiques n-aires, $n \geq 1$, il nous est nécessaire d'étudier, de façon générale, les relations rationnelles n-aires pour $n \geq 2$. Introduites dès 1965 par Elgot et Mezei [E.M] sous le nom de transductions n-aires, les relations rationnelles n-aires ont fait l'objet de nombreuses études dans le cas $n = 2$ sous le nom de transductions rationnelles. Les principaux résultats établis dans ce cas sont, sans aucun doute, la caractérisation en terme de bimorphismes alphabétiques donnée par Nivat [Ni1] et les nombreuses propriétés établies grâce à elle [Berst, Bo1, BoNi1, Beau2, Chof, Aute,...].

Dans le cas $n > 2$, Eilenberg [Ei2] donne un certain nombre de résultats sur les relations rationnelles n-aires et de nouvelles preuves de théorèmes de [E.M].

Les résultats que nous présentons dans ce chapitre s'obtiennent le plus souvent par des généralisations simples des preuves de [Ni1, Ei2, BoNi1] que nous ne donnerons donc pas. En fait ces résultats ne sont pas surprenants, nous les citons uniquement parce que, à notre connaissance, ils n'ont jamais été publiés.

Nous donnons tout d'abord une caractérisation en termes de morphismes des relations rationnelles n-aires analogue à celle de Nivat. Nous étudions ensuite des relations particulières utilisées dans le chapitre 5 et terminons par l'étude des applications rationnelles f de $X_1^* \times \ldots \times X_i^*$ dans $X_{i+1}^* \times \ldots \times X_n^*$ dont le graphe $\# f = \{(x_1,\ldots,x_n) \mid (x_{i+1},\ldots,x_n) \in f(x_1,\ldots,x_i)\}$ est une partie rationnelle. Nous généralisons alors des résultats de [Ei2] et [BoNi1]. Enfin nous constatons qu'un cône rationnel n'est pas nécessairement clos pour ces applications si $n > 2$.

B) DEFINITIONS ET PROPRIETES

Définition 4.1

Etant donné n alphabets finis $X_1,\ldots,X_n$, nous désignons par $Rat_n(X_1,\ldots,X_n)$ la famille des parties R de $X_1^* \times \ldots \times X_n^*$ qui peuvent se mettre sous la forme :

$$R = \{(\varphi_1 w,\ldots, \varphi_n w) \mid w \in K\} \qquad \text{avec}$$

- $K = AZ^* \cap Z^*B \setminus Z^*VZ^*$, langage rationnel local sur Z

$$(A \text{ et } B \subset Z,\ V \subset Z^2).$$

- $\varphi_1,\ldots, \varphi_n$ homomorphismes de Z^* dans $X_1^*,\ldots,X_n^*$ respectivement.

Tout élément de $Rat_n(X_1,\ldots,X_n)$ est appelé relation rationnelle n-aire sur $X_1,\ldots,X_n$.

Si $X_1 = \ldots = X_n = X$ nous noterons $Rat_n(X)$ l'ensemble $Rat_n(X,\ldots,X)$, et Rat(X) désignera $Rat_1(X)$.

De façon générale nous considèrerons trois opérations internes sur les relations n-aires et une opération externe :

Définition 4.2

Soient R et R' deux relations n-aires de $X_1^* \times \ldots \times X_n^*$.

Le produit R.R' est défini par

$$R.R' = \{(r_1r'_1,\ldots,r_nr'_n) \mid (r_1,\ldots,r_n) \in R, (r'_1,\ldots,r'_n) \in R'\}.$$

L'union $R \cup R'$ est définie comme l'union ensembliste.

L'étoile de R est :

$$R^* = \{(\varepsilon,\varepsilon)\} \cup \left(\bigcup_{k=1}^{\infty} R^k \right) \qquad \text{avec}$$

$$R^1 = R \quad \text{et} \quad \forall k \geq 1,\ R^{k+1} = R^k.R$$

Si $R \subset X_1^* \times \ldots \times X_n^*$ et $R' \subset X_n^* \times Y_1^* \times \ldots \times Y_m^*$, la composition de R par R' est la partie $X_1^* \times \ldots \times X_{n-1}^* \times Y_1^* \times \ldots \times Y_m^*$ définie par :

$$RoR' = \{(f_1,\ldots,f_{n-1},g_1,\ldots,g_m) \mid \exists f_n \in X_n^* : (f_1,\ldots,f_{n-1},f_n) \in R,\ (f_n,g_n,\ldots,g_m) \in R'\}.$$

Définition 4.3

Soit $\mathcal{S}$ une famille de parties de $X_1^* \times \ldots \times X_n^*$. La fermeture de Kleene de $\mathcal{S}$ est la plus petite famille $K(\mathcal{S})$ de parties de $X_1^* \times \ldots \times X_n^*$ telle que :

1) $\mathcal{S} \subseteq K(\mathcal{S})$

2) Si R et $R' \in K(\mathcal{S})$ alors $R \cup R'$, R.R' et R^* sont dans $K(\mathcal{S})$.

Nous avons alors le théorème :

Théorème 4.1

$Rat_n(X_1,\ldots,X_n)$ est identique à la fermeture de Kleene de la famill des parties finies de $X_1^* \times \ldots \times X_n^*$.

Ce théorème s'établit de façon analogue aux preuves de [Ei1] et [Ni1] à partir des lemmes suivants :

Lemme 4.1 [E.M]

Si R et R' sont deux éléments de $Rat_n(X_1,\ldots,X_n)$, alors $R \cup R'$ et $R.R'$ appartiennent à $Rat_n(X_1,\ldots,X_n)$.

Lemme 4.2 [Ei1]

Si $R \in Rat_n(X_1,\ldots,X_n)$ alors $R.R^* \in Rat_n(X_1,\ldots,X_n)$.

Lemme 4.3 [Ei1]

Toute partie finie de $X_1^* \times \ldots \times X_n^*$ est dans $Rat_n(X_1,\ldots,X_n)$.

Lemme 4.4

Si K est un langage rationnel quelconque sur Y et $\varphi_1,\ldots,\varphi_n$ n homomorphismes de Y^* dans $X_1^*,\ldots,X_n^*$ respectivement. Alors

$$R = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K\} \in Rat_n(X_1,\ldots,X_n).$$

Lemme 4.5

Soient K et K' deux langages rationnels sur Y^*, $\varphi_1,\ldots,\varphi_n$ n homomorphismes de Y^* sur $X_1^*,\ldots,X_n^*$ respectivement et $R = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K\}$, $R' = \{(\varphi_1 g,\ldots,\varphi_n g) \mid g \in K'\}$. Alors

$$R \cup R' = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K \cup K'\}$$

$$R.R' = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K.K'\}$$

$$R^* = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K^*\}.$$

Démonstration du théorème 4.1

Les lemmes 4.1, 2 et 3 impliquent l'inclusion de la fermeture de Kleene des parties finies de $X_1^* \times \ldots \times X_n^*$ dans $Rat_n(X_1,\ldots,X_n)$.

La réciproque résulte du théorème de Kleene pour les langages rationnels et de la remarque que si $K \subset Y^*$ est fini, alors pour tous homomorphismes $\varphi_1,\ldots,\varphi_n$ de Y^* dans $X_1^*,\ldots,X_n^*$ respectivement l'ensemble $\{(\varphi_1 h,\ldots,\varphi_n h) | h \in K\}$ est fini.

■

Exemples

1) $R_1 = \{(w,w,\ldots,w) | w \in K\}$, K langage rationnel quelconque sur X^* est dans $Rat_n(X)$.

2) $R_2 = \{(a^{p_1},a^{p_2},\ldots,a^{p_{n-1}},a^{p_1+\ldots+p_{n-1}}) | p_i \in \mathbb{N},\ 1 \leq i \leq n-1\}$ est dans $Rat_n(\{a\})$. Il suffit de vérifier que $R_2 = \{(\varphi_1 f,\ldots,\varphi_n f) | f \in \{a_1,\ldots,a_{n-1}\}^*\}$ avec $\varphi_i(a_j) = \begin{cases} a \text{ si } i = j \\ \varepsilon \text{ sinon} \end{cases}$, pour tout i, $1 \leq i \leq n-1$ et $\varphi_n(a_j) = a$, $\forall j$.

3) $R_3 = \{(a^{n_1},a^{n_2},a^{n_1 n_2}) | n_1,n_2 \geq 1\}$ n'est pas dans $Rat_3(\{a\})$. En effet de la caractérisation donnée de Rat_n il résulte immédiatement que si $R \in Rat_n$ alors, $\forall i$, $1 \leq i \leq n$, la ième projection de R :

$$\pi_i(R) = \{f_i | \exists f_1,\ldots,f_{i-1},f_{i+1},\ldots,f_n : (f_1,\ldots,f_n) \in R\}$$

est rationnelle.

Or $\pi_3(R_3) = \{a^{n_1 n_2} | n_1,n_2 > 1\}$ est l'ensemble des nombres non premiers qui est vérifiable [Rod1] ("Checking automata language" [Gr5]) mais

non rationnel.

Pour l'opération externe de composition rappelons le résultat d'Elgot et Mezei [E.M] :

Théorème 4.2 [E.M]

La composition de $R \in Rat_n(X_1,\ldots,X_n)$ par $R' \in Rat_{p+1}(X_n,Y_1,\ldots,Y_p)$ appartient à $Rat_{n+p-1}(X_1,\ldots,X_{n-1},Y_1,\ldots,Y_p)$.

Ce théorème résulte directement du lemme 4.6 suivant par un procédé analogue à celui de [Ni1] :

Lemme 4.6

Pour tout $R \in Rat_n(X_1,\ldots,X_n)$ il existe un langage rationnel local K sur Z^* et n homomorphismes alphabétiques ψ_i de Z^* dans X_i^*, $1 \leq i \leq n$, tels que :

$$R = \{(\varphi_1 h,\ldots,\varphi_n h) \mid h \in K\}.$$

Notation

Nous dirons qu'un ensemble $A \in Rat_n$ si et seulement si il existe n alphabets finis $X_1,\ldots,X_n$ tels que $A \in Rat_n(X_1,\ldots,X_n)$.

Proposition 4.1

Soit $R \in Rat_n$ et σ une permutation sur $\{1,\ldots,n\}$. La partie $R_\sigma = \{(f_{\sigma 1},\ldots,f_{\sigma n}) \mid (f_1,\ldots,f_n) \in R\}$ appartient à Rat_n.

Démonstration

Si $R \in Rat_n$, il existe un langage rationnel K et n homomorphismes φ_i, $1 \leq i \leq n$, tels que $R = \{(\varphi_1 f, \ldots, \varphi_n f) | f \in K\}$. Trivialement on a $R_\sigma = \{(\varphi_{\sigma 1} f, \ldots, \varphi_{\sigma n} f) | f \in K\}$, et donc $R_\sigma \in Rat_n$.

■

Proposition 4.2

Si $A_1 \in Rat_n(X_1, \ldots, X_n)$ et $A_2 \in Rat_m(Y_1, \ldots, Y_m)$ alors la partie $A_1 \times A_2 = \{(f_1, \ldots, f_n, g_1, \ldots, g_m) | (f_1, \ldots, f_n) \in A_1, (g_1, \ldots, g_m) \in A_2\}$ appartient à $Rat_{n+m}(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$.

Démonstration

D'après la définition de Rat_n, il existe deux alphabets finis Z et Z', deux langages rationnels $K \subseteq Z^*$ et $K' \subseteq Z'^*$, des homomorphismes φ_i de Z^* dans X_i^*, $1 \leq i \leq n$ et φ'_j de Z'^* dans Y_j^*, $1 \leq j \leq m$, tels que $A_1 = \{(\varphi_1 f, \ldots, \varphi_n f) | f \in K\}$ et $A_2 = \{(\varphi'_1 f, \ldots, \varphi'_m f)\ f \in K'\}$. On peut toujours supposer que $Z \cap Z' = \emptyset$.

Soit alors ψ_i l'homomorphisme de $(Z \cup Z')^*$ dans X_i^* défini par $\psi_i(a) = \varphi_i(a)$ si $a \in Z$, ε sinon, $1 \leq i \leq n$ et ψ'_j l'homomorphisme de $(Z \cup Z')^*$ dans Y_j^* défini par $\psi'_j(a) = \varphi'_j(a)$ si $a \in Z'$, ε sinon, $1 \leq j \leq m$.

On a alors clairement

$$A_1 \times A_2 = \{(\psi_1 f, \ldots, \psi_n f, \psi'_1 f, \ldots, \psi'_m f) | f \in KK'\},$$

i.e. $A_1 \times A_2 \in Rat_{n+m}(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$

■

Proposition 4.3

Soit X et X_i, $1 \leq i \leq n$ des alphabets finis et σ_i une substitution rationnelle de X^* dans X_i^*, $1 \leq i \leq n$. Alors

$$R = \{(x,x_1,\ldots,x_n) \mid x \in X^*,\ x_i \in \sigma_i(x),\ 1 \leq i \leq n\}$$

appartient à $Rat_{n+1}(X,X_1,\ldots,X_n)$.

Démonstration

$\forall x \in X$, l'ensemble $\{x\} x\ \sigma_1 x\ x \ldots x\ \sigma_n x$ est dans Rat_{n+1}, d'après la proposition précédente. Or nous avons :

$$R = \Big(\bigcup_{x \in X} \{x\} x\ \sigma_1 x\ x \ldots x\ \sigma_n x\Big)^*.$$ D'où le résultat.

■

Corollaire 4.1

Si φ_i est un homomorphisme de Y^* dans X_i^*, $1 \leq i \leq n$, alors quelque soit i, $1 \leq i \leq n$, la relation

$$R = \{(\varphi_1 w,\ldots,\varphi_{i-1} w, w, \varphi_{i+1} w,\ldots,\varphi_n w) \mid w \in X^*\}$$

appartient à Rat_n.

Démonstration

Ceci résulte directement de la proposition précédente et de la proposition 4.1.

■

Corollaire 4.2

Si $R_1 \in Rat_n(X_1,\ldots,X_n)$ et $R_2 \in Rat_n(Y_1,\ldots,Y_m)$ alors

$$R_{ij} = \{(f_1,\ldots,f_i,f'_1,\ldots,f'_j,f_{i+1},\ldots,f_n,f'_{j+1},\ldots,f'_m) \mid (f_1,\ldots,f_n) \in R$$
$$(f'_1,\ldots,f'_m) \in R_2)\}$$

appartient à $Rat_{n+m}(X_1,\ldots,X_i,Y_1,\ldots,Y_j,X_{i+1},\ldots,X_n,Y_{j+1},\ldots,Y_m)$.

Démonstration

Résulte de l'application successive des propositions 4.2 et 4.1.

Proposition 4.3

Pour tout $A \subseteq X_1^* \times\ldots\times X_n^*$, la relation

$$\cap A = \{(x_1,\ldots,x_n,x_1,\ldots,x_n) \mid (x_1,\ldots,x_n) \in A\} \in Rat_{2n}$$

si et seulement si $A \in Rat_n$.

Démonstration

Si $A \in Rat_n$ alors il existe un rationnel K et des homomorphismes $\varphi_1,\ldots,\varphi_n$ tels que $A = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K\}$. Or

$$\cap A = \{(\varphi_1 f,\ldots,\varphi_n f,\varphi_1 f,\ldots,\varphi_n f) \mid f \in K\}.$$

Donc $\cap A \in Rat_2$. La réciproque est triviale.

■

Corollaire 4.3

Si $X_1,\ldots,X_n$ sont des alphabets finis, alors la relation

$$R = \{(x_1,\ldots,x_n,x_1,\ldots,x_n) | x_i \in X_i^*,\ 1 \leq i \leq n\}$$

appartient à Rat_{2n}.

Proposition 4.5

Soient X et X_i, $1 \leq i \leq n$, des alphabets finis et $A_i \in Rat(X_i)$, $1 \leq i \leq n$. Alors la relation

$$R = \{(x,xa_1,\ldots,xa_n) | x \in X^*,\ a_i \in A_i,\ 1 \leq i \leq n\}$$

appartient à Rat_{n+1}.

Démonstration

On a de façon évidente :

$$R = (\{\varepsilon\} \times A_1 \times \ldots \times A_n).\{(x,\ldots,x) | x \in X^*\}.$$

■

De façon analogue à [Ei2], nous pouvons énoncer :

Théorème 4.3

Si $X_1,\ldots,X_n$ sont des alphabets finis disjoints deux à deux, une relation $R \subseteq X_1^* \times \ldots \times X_n^*$ est rationnelle si et seulement si il existe un langage rationnel $K \subseteq (X_1 \cup \ldots \cup X_n)^*$ tel que $R = \{(\pi_1 x,\ldots,\pi_n x) | x \in K\}$ avec $\pi_i(a) = a$ si $a \in X_i$, ε sinon.

Démonstration

Si R est de la forme ci-dessus alors R appartient à Rat_n de façon triviale.

Réciproquement, si $R \in Rat_n(X_1,\ldots,X_n)$ comme le morphisme π de $(X_1 \cup \ldots \cup X_n)^*$ dans $X_1^* \times \ldots \times X_n^*$ défini par $\pi y = (\pi_1 y,\ldots,\pi_n y)$ est surjectif, il en résulte que $\pi(Rat(X_1 \cup \ldots \cup X_n)) = Rat(X_1^* \times \ldots \times X_n^*)$ et donc il existe $A \in (X_1 \cup \ldots \cup X_n)^*$ tel que $R = \pi A$.

■

C) LEMME D'ITERATION

Du théorème 4.3 et du lemme du l'étoile nous pouvons déduire :

Lemme 4.7

Soit $R \in Rat_n(X_1,\ldots,X_n)$. Alors il existe un entier k tel que si $(f_1,\ldots,f_n) \in R$ et k lettres ou plus de $f_1,\ldots,f_n$ sont considérées comme distinguées, alors chaque f_i a une factorisation $f_i = x_i y_i z_i$, $1 \leq i \leq n$, telle que le mot $y_1 y_2 \ldots y_n$ contient au moins une position distinguée et pour tout $p \geq 0$, $(x_1 y_1^p z_1,\ldots,x_n y_n^p z_n) \in R$.

Démonstration

Sans nuire à la généralité nous pouvons supposer les alphabets X_i, $1 \leq i \leq n$, disjoints deux à deux. Alors d'après le théorème 4.3, il existe un langage rationnel $A \subseteq (X_1 \cup \ldots \cup X_n)^*$ tel que $R = \{(\pi_1 f,\ldots,\pi_n f) \mid f \in A\}$. Soit alors k l'entier associé à A par le lemme de l'étoile. Si k ou plus positions de $f_1 \ldots f_n$ sont considérées comme distinguées et si $w \in A$ est tel que $(f_1,\ldots,f_n) = (\pi_1 w,\ldots,\pi_n w)$, alors il existe x, y et $z \in (\bigcup_{i=1}^{n} X_i)^*$ tel que y contient une position distinguée au moins, $w = xyz$ et $xy^*z \subseteq A$.

Soit alors $x_i = \pi_i x$, $y_i = \pi_i y$ et $z_i = \pi_i z$, $1 \leq i \leq n$, on obtient la factorisation énoncée.

■

Remarque

La réciproque du lemme 4.7 est fausse. Ainsi, par exemple la relation $R = \{(a^n,b^p) \mid n,p \text{ non premiers}\}$ n'est pas rationnelle, puisque chacune de ses projections n'est pas un langage rationnel. Cependant pour tout n et $p \geq 2$, n et p non premiers $(a^{nq},b^{qp}) \in R$, $\forall q \geq 0$.

D) RELATIONS RATIONNELLES PARTICULIERES

Proposition 4.6

Si $K_1,\ldots,K_n$ sont n langages rationnels sur $X_1,\ldots,X_n$ respectivement, alors $K_1 \times\ldots\times K_n \in Rat_n(X_1,\ldots,X_n)$.

Démonstration

On a $K_1 \times\ldots\times K_n = (K_1 \times \{\varepsilon\} \times\ldots\times \{\varepsilon\})\ldots(\{\varepsilon\} \times\ldots\times \{\varepsilon\} \times K_n)$ et chacun des facteurs $\{\varepsilon\} \times\ldots\times \{\varepsilon\} \times K_i \times \{\varepsilon\} \times\ldots\times \{\varepsilon\}$ appartient à $Rat_n(X_1,\ldots,X_n)$.

∎

Nous allons maintenant étudier la relation produit (ou concaténation).

Lemme 4.8

Soient $X_1,\ldots,X_n$ des alphabets disjoints deux à deux et $R \in Rat_n(X_1,\ldots,X_n)$. Une condition suffisante pour que la relation $R' = \{(f_1,\ldots,f_n,f_1\ldots f_n) \mid (f_1,\ldots,f_n) \in R'\}$ appartienne à Rat_{n+1} est qu'il existe des langages rationnels $A_i \subseteq X_i^*$, $1 \leq i \leq n$ tel que $R = A_1 \times\ldots\times A_n$.

Démonstration

Si $A_i \in Rat(X_i)$, $1 \leq i \leq n$, alors le langage $A_1.A_2\ldots A_n$ est rationnel. On a alors clairement $R = \{(\pi_1 f,\ldots,\pi_n f) \mid f \in A_1\ldots A_n\}$, avec $\pi_i(a) = a$ si $a \in X_i$, ε sinon. Il en résulte que $R' = \{(\pi_1 f,\ldots,\pi_n f,f) \mid f \in A_1\ldots A_n\}$ et donc $R' \in Rat_{n+1}$ d'après le corollaire 4.1.

∎

Lemme 4.9

Soit X_i, $1 \leq i \leq n$, des alphabets finis disjoints deux à deux. Si la relation $R = \{(f_1, f_2, \ldots, f_n, f_1 \ldots f_n)\}$ appartient à $Rat(X_1, \ldots, X_n, \bigcup_{i=1}^{n} X_i)$ alors la relation $R' = \{(f_1, \ldots, f_n) | (f_1, \ldots, f_n, f_1 \ldots f_n) \in R\}$ appartient à $Rat_n(X_1, \ldots, X_n)$ et il existe $k \in \mathbb{N}$ tel que si $(f_1, \ldots, f_n) \in R'$ et k lettres ou plus de f_i sont considérées comme distinguées, $1 \leq i \leq n$, alors il existe x,y et $z \in X_i^*$ tels que y contient au moins une position distinguée et $(f_1, \ldots, f_{i-1}, xy^*z, f_{i+1}, \ldots, f_n) \subseteq R'$.

Démonstration

Si $R \in Rat_{n+1}$, alors, trivialement $R' \in Rat_n$. Etant donné la disjonction des alphabets X_i, $1 \leq i \leq n$, $R \in Rat_{n+1}$ si et seulement si la relation $R_1 = \{(f_1, \ldots, f_n, \overline{f_1} \ldots \overline{f_n}) | (f_1, \ldots, f_n, f_1 \ldots f_n) \in R\}$ appartient à Rat_{n+1}, où $\overline{f_i}$ est une recopie de $f_i \in X_i^*$ sur un alphabet $\overline{X}_i$ tel que les alphabets X_i et $\overline{X}_i$, $1 \leq i \leq n$, soient disjoints deux à deux.

D'après le théorème 4.3, il existe donc un langage rationnel $A \subseteq (\bigcup_{i=1}^{n} (X_i \cup \overline{X}_i))^*$ tel que $R_1 = \{(\pi_1 f, \ldots, \pi_n f, \pi_{n+1} f) | f \in A\}$. Soit k l'entier associé à A par le lemme de l'étoile. Si $(f_1, \ldots, f_n, \overline{f_1} \ldots \overline{f_n}) \in R_1$ et k lettres au moins de f_i sont considérées comme distinguées, alors le mot $w \in A$ tel que $(f_1, \ldots, f_n, \overline{f_1} \ldots \overline{f_n}) = (\pi_1 w, \ldots, \pi_{n+1} w)$ admet une factorisation $w = xyz$ telle que y contient au moins une position distinguée et $xy^*z \subseteq A$.

Il en résulte que $(\pi_1 xy^*z, \ldots, \pi_{n+1} xy^*z) \subseteq R_1$ et donc y contient autant d'occurrences de toute lettre $x_k \in X_k$ que d'occurrence de $\overline{x}_k$, $1 \leq k \leq n$.

Montrons alors que $y \in (X_i \cup \overline{X}_i)^*$. En effet si y contient une occurrence de $x_i \in X_i$ -et y en contient au moins une d'après le lemme de l'étoile- et une occurrence de $x_j \in X_j$, $i \neq j$, alors y^2 contient un facteur de la forme $\alpha\overline{x}_i\ \beta\overline{x}_j\ \gamma\overline{x}_i\ \delta\overline{x}_j\ \mu$ ou bien de la forme $\alpha\overline{x}_j\ \beta\overline{x}_i\ \gamma\overline{x}_j\ \delta\overline{x}_i\ \mu$ et dans les deux cas $\pi_{n+1}\ xy^2z$ n'est pas inclus dans $X_1^*\ldots X_n^*$, ce qui est contradictoire.

Il en résulte donc que $\pi_j(xy^pz) = \pi_j(xyz)$, $\forall p \geq 0$ et $j \neq i$, autrement dit :

$$(\pi_1 xy^*z,\ldots,\pi_{i-1}xy^*z,\ \pi_i xy^*z,\ \pi_{i+1}xy^*z,\ldots,\ \pi_{n+1}xy^*z) =$$

$$(f_1,\ldots,f_{i-1},\ \alpha\beta^*\gamma,\ f_{i+1},\ldots f_{n+1}) \subseteq R_1 \text{ si } \pi x = \alpha,\ \pi y = \beta \text{ et}$$

$$\pi z = \gamma \text{ et donc } (f_1,\ldots,f_{i-1},\ \alpha\beta^*\gamma,\ f_{i+1},\ldots,f_n) \subseteq R'.$$

■

Exemples

Il résulte directement des deux lemmes précédents que :

- $R_1 = \{(a^n,b^m,c^{n+m}) \mid n,m \in \mathbb{N}\} \in Rat_3$.

- $R_2 = \{(a^n,b^n,a^nb^n) \mid n \in \mathbb{N}\} \notin Rat_3$.

- $R_3 = \{(a^n,b^m,c^{n+m},a^nb^mc^{n+m}) \mid n,m \in \mathbb{N}\} \notin Rat_4$.

Le problème reste ouvert de savoir si les conditions des deux lemmes précédents sont nécessaires et suffisantes pour que, si $R \in Rat_n$ alors $R' = \{(f_1,\ldots,f_n,f_1\ldots f_n) \mid (f_1,\ldots,f_n) \in R\}$ appartienne à Rat_{n+1}.

Nous pouvons seulement énoncer :

Lemme 4.10

Soient X_i, $1 \leq i \leq n$, des alphabets finis disjoints deux à deux. Si la relation $R \in Rat_n(X_1,\ldots,X_n)$ est telle qu'il existe $k \in \mathbb{N}$, tel que si $(f_1,\ldots,f_n) \in R$ et k occurrences de lettres au moins de f_i sont considérées comme distinguées, alors il existe une factorisation de $f_i = xyz$, telle que y contient au moins une position distinguée et

$$\pi_1 R \times \ldots \times \pi_{i-1} R \times (xy^*z) \times \pi_{i+1} R \times \ldots \times \pi_n R \subseteq R$$

alors il existe des langages rationnels A_i, $1 \leq i \leq n$, tels que $R = A_1 \times A_2 \times \ldots \times A_n$.

Démonstration

Soit $R_f = \{(f_1,\ldots,f_n) \notin R \mid |f_i| < k, 1 \leq i \leq n\}$. R_f est fini, donc le langage $A = \pi_1 R \ldots \pi_n R \backslash \{f_1 \ldots f_n \mid (f_1,\ldots,f_n) \in R_f\}$ est rationnel puisque $R \in Rat_n$.

On a clairement $R \subseteq \pi_1 A \times \ldots \times \pi_n A$ si $\pi_i(a) = a$ si $a \in X_i$, ε sinon.

Réciproquement soit $(f_1,\ldots,f_n) \in \pi_1 A \times \ldots \times \pi_n A$. Deux cas sont possibles :

1) Si $|f_i| < k$ pour tout i, $1 \leq i \leq n$, alors $(f_1,\ldots,f_n) \in R$.

2) Il existe i tel que $|f_i| \geq k$, i.e. toutes les occurrences de lettres de f_i sont considérées comme distinguées. Alors comme $f_i \in \pi_i R$, il existe des mots $f'_1,\ldots,f'_{i-1},f'_{i+1},\ldots,f'_n$ et des mots x,y,z tels que $f_i = xyz$ et $(f'_1,\ldots,f'_{i-1},f_i,f'_{i+1},\ldots,f'_n) \in R$. D'après notre hypothèse on a donc $\pi_1 R \times \ldots \times \pi_{i-1} R \times \{xy^*z\} \times \pi_{i+1} R \times \ldots \times \pi_n R \subseteq R$ et donc en particulier

$(f_1,\ldots,f_i,\ldots,f_n) \in R$, ce qui achève la démonstration, si l'on prend $A_i = \pi_i A$, $1 \leq i \leq n$.

Nous terminerons ce paragraphe par l'étude de relations rationnelles plus particulières :

Proposition 4.7

Si $\varphi_1,\ldots, \varphi_n$ sont n homomorphismes de $X_1^*,\ldots,X_n^*$ dans Y^*, la relation :

$$R = \{(f_1,\ldots,f_n) \mid \varphi_1 f_1 = \ldots = \varphi_n f_n\} \in \mathrm{Rat}_n(X_1,\ldots,X_n)$$

Démonstration

Supposons $\varphi_1,\ldots, \varphi_n$ alphabétiques. Soient alors les ensembles finis :

$$Z_0 = \{(x_1,\ldots,x_n) \mid x_i \in X_i,\ \varphi_1(x_1) = \ldots = \varphi_n(x_n)\}$$

et si c est une lettre, $c \notin \bigcup_i X_i$:

$$Z_1 = \{(x_1,c,\ldots,c) \mid x_1 \in X_1,\ \varphi_1 x_1 = \varepsilon\}$$

$$\ldots$$

$$Z_n = \{(c,\ldots,c,x_n) \mid x_n \in X_n,\ \varphi_n x_n = \varepsilon\}.$$

Soient alors φ_i' un homomorphisme de $Z^* = (Z_0 \cup \ldots \cup Z_n)^*$ dans X_i^*, $1 \leq i \leq n$, défini par :

$$\varphi'_i(x_1,\ldots,x_n) = x_i, \quad \varphi'_i(c,\ldots,c,x_i,c,\ldots,c) = x_i$$

$$\varphi'_i(c,\ldots,c,x_j,c,\ldots,c) = \varepsilon \quad \text{si} \quad i \neq j.$$

Montrons alors que :

$$R = \{(\varphi'_1 g,\ldots,\varphi'_n g) \mid g \in Z^*\}.$$

En effet pour tout i et j :

$$\varphi_i \varphi'_i(x_1,\ldots,x_n) = \varphi_i x_i = \varphi_j x_j = \varphi_j \varphi'_j(x_1,\ldots,x_n)$$

et $$\varphi_i \varphi'_i(c,\ldots,c,x_i,c,\ldots,c) = \varphi_i x_i = \varepsilon = \varphi_j \varepsilon = \varphi_j \varphi'_j(c,\ldots,c,x_i,c,\ldots c)$$

Donc pour tout $g \in Z^*$, $\varphi_i \varphi'_i g = \varphi_j \varphi'_j g$,
i.e. $\{(\varphi'_1 g,\ldots,\varphi'_n g) \mid g \in Z^*\} \subseteq R$.

Réciproquement soient n mots $f_1,\ldots,f_n$ tels que $\varphi_1 f_1 = \ldots = \varphi_n f_n$ avec $f_i = x_{i-1}\ldots x_{ik}$. Il existe donc pour tout i, ℓ symboles de la suite $x_{i1},\ldots,x_{ik}$ dont l'image par φ_i est différente de ε. Soit $y_{i1},\ldots,y_{i\ell}$ ces suites. Il est clair que

$$\varphi_i(y_{ip}) = \varphi_j(y_{jp}), \ \forall i,j,p.$$

Soit alors h le mot de Z^* :

$$h = (y_{11},y_{21},\ldots,y_{n1})\ldots(y_{1\ell},y_{2\ell},\ldots,y_{n\ell})$$

et soit h' le mot de Z^* obtenu à partir de h en intercalant entre les lettres de $h \in Z$, les n-uplets comportant au moint une occurrence de "c", afin que

pour tout i :

$$\varphi_i' h' = f_i$$

Plus précisément :

$$h' = (x_{11},c,\ldots,c)(x_{12},c,\ldots,c)\ldots(x_{1p},c,\ldots,c)(c,x_{21},c,\ldots,c)\ldots$$
$$(c,\ldots,c,x_{nr})(y_{11},y_{21},\ldots,y_{n1})\ldots(y_{1\ell},\ldots,y_{n\ell}).$$

Il en résulte que

$$R \subseteq \{(\varphi'_{1g},\ldots,\varphi'_{ng}) \mid g \in Z^*\}.$$

La proposition résulte alors du lemme 4.6.

■

Proposition 4.8

Si $\varphi_1,\ldots,\varphi_{n-1}$ sont des homomorphismes de Y^* dans T^* et φ_n un homomorphisme de X^* dans T^*, alors :

$$R_1 = \{(f_1,\ldots,f_n) \mid \varphi_n f_1 = \varphi_1 f_n,\ \varphi_n f_2 = \varphi_2 f_n,\ldots,\ \varphi_n f_{n-1} = \varphi_{n-1} f_n\}$$

appartient à $Rat_n(X,\ldots,,Y)$.

Démonstration

Le procédé utilisé est identique à celui de la démonstration de la proposition 4.5.

On peut supposer que $\varphi_1,\ldots,\varphi_n$ sont alphabétiques. Soient alors les ensembles finis :

$$Z_0 = \{(x_1,\ldots,x_n) \mid \varphi_n x_1 = \varphi_1 x_n, \ldots, \varphi_n x_{n-1} = \varphi_{n-1} x_n, x_i \in X, \quad 1 \leq i \leq n-1,\ x_n \in Y\}.$$

$\forall i,\ 1 \leq i \leq n-1,\ Z_i = \{(c,\ldots,c,x_i,c,\ldots,c) \mid x_i \in X,\ \varphi_n(x_i) = \varepsilon\}$

et $\quad Z_n = \{(c,\ldots,c,x_n) \mid \varphi_i(x_n) = \varepsilon,\ 1 \leq i \leq n-1, x_n \in Y\}.$

Si l'on définit alors $\forall i,\ 1 \leq i \leq n$:

$$\varphi_i'(x_1,\ldots,x_n) = x_i, \quad \varphi_i'(c,\ldots,c,x_i,c,\ldots,c) = x_i$$

$$\varphi_i'(c,\ldots,c,x_j,c,\ldots,c) = \varepsilon \quad \text{si} \quad i \neq j$$

alors par un procédé analogue à celui utilisé pour la proposition 4.7, on montre que $R_1 = \{(\varphi_1' g,\ldots,\varphi_n' g) \mid g \in Z^*\}$ si $Z = \bigcup_{i=0}^{n} Z_i$.

■

Proposition 4.9

Soit $R \in \mathrm{Rat}_n(X_1,\ldots,X_n)$, w et x deux mots appartenant à $X^* = (\bigcup_{i=1}^{n} X_i)^*$. Alors pour tout i, $2 \leq i \leq n$ l'ensemble :

$$R_i = \{(a_i,\ldots,a_n) \mid \exists a_1,\ldots,a_{i-1},w_1,\ldots,w_{i-2} \in X^* :$$

$$(a_1,\ldots,a_n) \in R,\ a_1 w_1 \ldots a_{i-2} w_{i-2} a_{i-1} w = x\}$$

appartient à $\mathrm{Rat}_{n-i+1}(X_i,\ldots,X_n)$.

Démonstration

Soit $K \subseteq Z^*$ un langage rationnel local et φ_i, $1 \leq i \leq n$, un homomorphisme alphabétique de Z^* dans X_i^* tels que :

$$R = \{(\varphi_1 f, \ldots, \varphi_n f) \mid f \in K\}.$$

Pour x et w donnés, la relation :

$$A = \{(a_1,\ldots,a_{i-1}) \mid \exists w_1,\ldots,w_{i-2} \in X^*, a_j \in X_j,\ 1 \leq j \leq i-1 :$$

$$a_1 w_1 \ldots a_{i-2} w_{i-2} a_{i-1} w = x\}$$

est finie donc $A \in \mathrm{Rat}_{i-1}(X_1,\ldots,X_{i-1})$. Il existe donc un langage rationnel $K' \subseteq T^*$ et i-1 homomorphismes $\psi_1,\ldots,\psi_{i-1}$ de T^* dans $X_1^*,\ldots,X_{i-1}^*$ respectivement, tels que

$$A = \{(\psi_1 f,\ldots,\psi_{i-1} f) \mid f \in K'\}.$$

On a alors clairement :

$$R_i = \{(\varphi_i f,\ldots,\varphi_n f) \mid f \in K \cap (\bigcap_{j=1}^{i-1} \psi_j^{-1} K') \text{ et } K \cap (\bigcap_{j=1}^{i-1} \psi_j^{-1} K')\}$$

étant rationnel, le résultat en découle.

■

Proposition 4.10

Soit $R \in \mathrm{Rat}_n(X_1,\ldots,X_n)$ et $w \in X_1^*$. Alors la relation :

$$w/R = \{(a_1', a_2, \ldots, a_n) \mid (wa_1', a_2, \ldots, a_n) \in R\}$$

appartient à $\mathrm{Rat}_n(X_1,\ldots,X_n)$.

Démonstration

$R = \{(\varphi_1 f, \ldots, \varphi_n f) \mid f \in K\}$, K rationnel $\subseteq Z^*$. Soit alors $c \notin Z$ et φ l'homomorphisme de $(Z \cup \{c\})^*$ dans Z^* défini par $\varphi(c) = \varepsilon$ et $\varphi(a) = a$, $\forall a \in Z$.

Soit $K' = \varphi^{-1}(K) \cap \varphi_1^{-1}(w) \subset Z^*$. K' est un langage rationnel. Soit alors τ la transduction rationnelle qui recopie les préfixes des mots de Z^* sur un alphabet T disjoint de $Z \cup \{c\}$ et laisse les suffixes de ces mots inchangés sitôt que la lettre "c" a été rencontrée et effacée. Soit $K_1 = Z(K')$, K_1 est rationnel. Soient alors n homomorphismes ψ_i, $1 \leq i \leq n$, de $(T \cup Z)^*$ dans X_i^* définis par :

$$\psi_1(a) = \varepsilon \qquad \forall a \in T,\ \psi_1(a) = a \text{ si } a \in Z$$

et pour tout $i \geq 2$:

$$\psi_i(a) = \overline{a} \qquad \forall a \in T, \text{ si } \overline{a} \text{ correspond à } a \in Z$$

$$\psi_i(a) = a \qquad \forall a \in Z.$$

Il est facile de vérifier alors que :

$w/R = \{(\psi_1 f, \ldots, \psi_n f) \mid f \in K_1\}$, i.e. $w/R \in \mathrm{Rat}_n(X_1, \ldots, X_n)$.

■

Proposition 4.11

Si $R \in \mathrm{Rat}_n$ et si φ est un homomorphisme alphabétique alors

$$\varphi^{-1}(R) = \{(\varphi^{-1}(a_1), \ldots, \varphi^{-1}(a_n)) \mid (a_1, \ldots, a_n) \in R\}$$

appartient à Rat_n.

Démonstration

Soit $R \in Rat_n(X_1,\ldots,X_n)$ et φ un homomorphisme alphabétique de T^* dans $(\bigcup_{i=1}^{n} X_i)^*$. Or

$$\varphi^{-1}(R) = \{(b_1,\ldots,b_n) \mid (\varphi b_1,\ldots,\varphi b_n) \in R\}$$

et

$$R = \{(\varphi_1 f,\ldots,\varphi_n f) \mid f \in K\}$$

pour K rationnel $\subseteq Z^*$ et φ_i homomorphisme alphabétique de Z^* dans X_i^*, $1 \leq i \leq n$. Soient alors

$$R_1 = \{(a_1,\ldots,a_{n-1},b) \mid \varphi(a_1) = \varphi_1(b),\ldots,\varphi(a_{n-1}) = \varphi_{n-1}(b)\}$$

$$R_2 = \{(b,b) \mid b \in K\}$$

$$R_3 = \{(b,a_n) \mid \varphi(a_n) = \varphi_n(b)\}$$

D'après la proposition 4.8, $R_1 \in Rat_n$ et $R_3 \in Rat_2$ et d'après le lemme 4.4, $R_2 \in Rat_2$.

Or $\varphi^{-1}(R) = R_1 o R_2 o R_3$. Donc $\varphi^{-1}(R) \in Rat_n$.

■

E) RELATIONS RATIONNELLES COMME APPLICATIONS

A toute relation rationnelle nous allons faire correspondre une famille finie d'applications, de façon analogue à la démarche de Nivat [Ni1] pour les transductions rationnelles binaires. La plupart des résultats donnés par Eilenberg [Ei2] ou Boasson et Nivat [NiBo1] se généralisent alors assez simplement.

Définition 4.4

Etant donnés n alphabets finis $X_1,\ldots,X_n$, une application $f : X_1^* x \ldots x X_i^* \to X_{i+1}^* x \ldots x X_n^*$ est rationnelle si et seulement si le graphe de f (noté $\neq f$) défini par

$\neq f = \{(f_1,\ldots,f_n) \in X_1^* x \ldots x X_n^* \mid (f_{i+1},\ldots,f_n) \in f(f_1,\ldots,f_i)\}$ appartient à $Rat_n(X_1,\ldots,X_n)$.

Ainsi à toute relation rationnelle $R \in Rat_n$ nous pourrons associer (n-1) applications rationnelles qui auront R comme graphe.

De la caractérisation de Rat_n il résulte immédiatement :

Théorème 4.4

Une application $f : X_1^* x \ldots x X_i^* \to X_{i+1}^* x \ldots x X_n^*$ est rationnelle si et seulement si il existe un alphabet fini Z, $K \subseteq Z^*$ un langage rationnel (local), $\varphi_1,\ldots,\varphi_n$ des homomorphismes (alphabétiques) de Z^* dans $X_1^*,\ldots,X_n^*$ respectivement tels que

$$f(x_1,\ldots,x_i) = \{(\varphi_{i+1}f,\ldots,\varphi_n f) \mid f \in \varphi_1^{-1}x_1 \cap \ldots \cap \varphi_i^{-1}x_i \cap A\}.$$

Suivant Boasson et Nivat [BoNi1] si l'on définit

Définition 4.5

Une application rationnelle f de $X_1^* \times \ldots \times X_i^* \to X_{i+1}^* \times \ldots \times X_n^*$ est :

- d'image finie si et seulement si

$\forall (x_1,\ldots,x_i) \in X_1^* \times \ldots \times X_i^*$, $f(x_1,\ldots,x_i)$ est une partie finie de $X_{i+1}^* \times \ldots \times X_n^*$;

- fidèle si et seulement si l'application rationnelle f^{-1} définie par $f^{-1}(x_{i+1},\ldots,x_n) = \{(x_1,\ldots,x_i) | (x_1,\ldots,x_i,\ldots,x_n) \in \not\asymp f\}$ est d'image finie ;

- continue si et seulement si $(\varepsilon,\ldots,\varepsilon) \in f(x_1,\ldots,x_i)$ implique $x_j = \varepsilon$, $\forall j$, $1 \leq j \leq i$.

Nous dirons que $\not\asymp f$ est respectivement d'image finie, fidèle ou continue.

On établit alors de façon analogue à [BoNi1] :

Proposition 4.12

L'application f de $X_1^* \times \ldots \times X_i^* \to X_{i+1}^* \times \ldots \times X_n^*$ est rationnelle fidèle si et seulement si il existe un alphabet Z, un langage rationnel R sur Z^* et des homomorphismes alphabétiques $\varphi_1,\ldots,\varphi_n$ de Z^* dans $X_1^*,\ldots,X_n^*$ respectivement tels que

1) $\not\asymp f = \{(\varphi_1 x,\ldots,\varphi_n x) | x \in R\}$

2) Pour tout mot $h \in R$, le nombre de lettres consécutives d'image vide par $\varphi_{i+1},\ldots,\varphi_n$ est borné.

Proposition 4.13

L'application $f : X_1^* x...x X_i^* \to X_{i+1}^* x...x X_n^*$ est rationnelle, fidèle et continue si et seulement si il existe un alphabet Z, un langage rationnel $K \subseteq Z^*$ des homomorphismes $\varphi_1,..., \varphi_i$ de Z^* dans $X_1^*,...,X_i^*$ respectivement et des homomorphismes strictement alphabétiques $\varphi_{i+1},..., \varphi_n$ de Z^* dans $X_{i+1}^*,...,X_n^*$ respectivement tels que

$$\neq f = \{(\varphi_1 x,..., \varphi_n x) | x \in K\}.$$

Remarque

On ne peut plus dans ce cas imposer aux homomorphismes φ_ℓ, $1 \leq \ell \leq i$ d'être alphabétiques [BoNi1].

Théorème 4.5

Soit $R \in Rat_n(X_1,...,X_n)$. Alors pour tout i, $1 \leq i < n$, si $A_1,...,A_i$ sont des parties rationnelles de $X_1^*,...,X_i^*$ respectivement, si $f_i : X_1^* x...x X_i^* \to X_{i+1}^* x...x X_n^*$ est une application rationnelle de graphe R, $f_i(A_1,...,A_i) \in Rat_{n-i}(X_{i+1},...,X_n)$.

Démonstration

Soit $R = \{(\varphi_1 f,..., \varphi_n f) | f \in K\}$, alors on a de façon triviale :

$$f_i(A_1,...,A_i) = \{(\varphi_{i+1} f,..., \varphi_n f) | f \in K \cap (\bigcap_{\ell=1}^{i} \varphi_\ell^{-1} A_\ell)\}.$$

$\varphi_\ell^{-1}(A_\ell)$ étant un langage rationnel, le résultat en découle immédiatement.

■

Remarque

Le résultat n'est pas vrai si l'on considère $f_i(A)$ pour $A \in Rat_i$, $i > 1$. En effet soit $A = \{(a^n b^p, c^p) | n,p \in \mathbb{N}\}$ et $B = \{(a^n b^p, c^n) | n \in \mathbb{N}\}$. A et $B \in Rat_2$ de façon triviale. Soit alors f_2 l'application rationnelle associée à la relation rationnelle $\cap A = \{(a^n b^p, c^p, a^n b^p, c^p) | n,p \in \mathbb{N}\}$. On a trivialement $f_2(B) = A \cap B = \{(a^n b^n, c^n) | n \in \mathbb{N}\}$ qui n'est pas une relation rationnelle.

Proposition 4.14

Soit $f : X_1^* x...x X_i^* \to X_{i+1}^* x...x X_n^*$ une application rationnelle. Alors pour tous langages rationnels $A_1,...,A_k$, $k < i$, l'application $g : X_{k+1}^* x...x X_i^* \to X_{i+1}^* x...x X_n^*$ définie par :

$$g(x_{k+1},...,x_i) = \{f(x_1,...,x_k,x_{k+1},...,x_i) | x_\ell \in A_\ell,\ 1 \leq \ell \leq k\}$$

est rationnelle.

Démonstration

Par hypothèse $\neq f \in Rat_n(X_1,...,X_n)$. Soit h l'application de $X_1^* x...x X_k^* \to X_{k+1}^* x...x X_n^*$ dont le graphe coïncide avec $\neq f$. h est donc une application rationnelle. D'après le théorème précédent $h(A_1,...,A_k) \in Rat_{n-k}(X_{k+1},...,X_n)$. Or $h(A_1,...,A_k)$ n'est autre que le graphe de g. Donc g est rationnelle.

∎

Définition 4.6

Etant donnée une relation $f(x_1,\ldots,x_n)$ et $\sigma = (i_1,\ldots,i_k)$ un sous-ensemble de $\{1,\ldots,n\}$. Nous désignerons par

$$f_\sigma^{-1}(x_{j_1},\ldots,x_{j_{n-k}}) = \{(x_{i_1},\ldots,x_{i_k}) | f(x_{j_1},\ldots,x_{i_1},\ldots,x_{i_2},\ldots,x_{i_k},\ldots,x_{j_{n-k}})\}.$$

Par extension nous désignerons par

$$f_\sigma^{-1}(A_{j_1},\ldots,A_{j_{n-k}}) = \{(x_{i_1},\ldots,x_{i_k}) | f(x_{j_1},\ldots,x_{i_1},\ldots,x_{i_k},\ldots,x_{j_{n-k}}) | x_{j_\ell} \in A_{j_\ell},$$
$$1 \leq \ell \leq n-k\}.$$

Proposition 4.15

Soit $R \in Rat_n(X_1,\ldots,X_n)$. Alors pour tout sous-ensemble $\sigma = (i_1,\ldots,i_k)$ de $\{1,n\}$, la relation R_σ^{-1} est rationnelle.

Démonstration

Si $R = \{(\varphi_1 f,\ldots,\varphi_n f) | f \in K\}$ alors on a trivialement $R_\sigma^{-1} = \{(\varphi_{i_1} f,\ldots,\varphi_{i_k} f) | f \in K\}$, ce qui démontre le résultat.

■

Proposition 4.16

Soit $R \in Rat_n(X_1,\ldots,X_n)$. Si $A_{j\ell} \in Rat(X_{j\ell})$, $1 \leq \ell \leq n-k$ alors $R_\sigma^{-1}(A_{j_1},\ldots,A_{j_{n-k}}) \in Rat_{n-k}$.

Démonstration

Résulte directement des deux propositions précédentes.

■

Corollaire 4.4

Si $A_1,\ldots,A_n$ sont des langages rationnels respectivement inclus dans $X_1^*,\ldots,X_n^*$, alors il en est de même de $A_1\ldots A_n$ et $P_\sigma^{-1}(A_{i_1},\ldots,A_{i_k})$, où $P(x_1,\ldots,x_n) = \{x_1\ldots x_n\}$.

Remarque

Par analogie avec [Ni1] considérons pour toute relation rationnelle n-aire R l'application f de X_1^* x...x X_{n-1}^* dans X_n^* définie par

$$f_R(x_1,\ldots,x_{n-1}) = \{x \mid (x_1,\ldots,x_{n-1},x) \in R\}.$$

Il est clair que toute famille de langages $\mathcal{L}$ fermée par ces applications est fermée par transduction rationnelle. La réciproque est fausse. Considérons en effet $\mathcal{L}$ le cône rationnel engendré par le langage $L_= = \{a^n b^n \mid n \in \mathbb{N}\}$. Soient alors φ_1 et φ_2 les homomorphismes définis par $\varphi_1(\alpha) = \varepsilon$, $\varphi_1(\beta) = a$, $\varphi_1(\gamma) = b$ et $\varphi_2(\alpha) = a$, $\varphi_2(\beta) = b$ et $\varphi_2(\gamma) = \varepsilon$ et soit $K = \alpha^* \beta^* \gamma^*$. Si l'on considère la relation rationnelle ternaire $R = \{(\varphi_1(x), \varphi_2(x), x) \mid x \in K\}$ et f_R l'application associée, on a alors d'après le théorème 4.4 :

$$f_R(L_=,L_=) = \{w \in \varphi_1^{-1}(L_=) \cap \varphi_2^{-1}(L_=) \cap K\}$$

soit $$f_R(L_=,L_=) = \{\alpha^n \beta^n \gamma^n \mid n \in \mathbb{N}\}$$

et ce langage n'appartient pas au cône rationnel engendré par $L_=$.

Ainsi les langages algébriques ne sont pas fermés par "transduction rationnelle n-aire", n > 2.

CHAPITRE 5

CHEVRON ET CROCHET OUVERT D'ORDRE N

A) INTRODUCTION

B) PROPRIETES GENERALES

C) APPLICATIONS

A) INTRODUCTION

Diverses généralisations unaires du crochet, chevron et crochet ouvert, différentes de celles du chapitre 3, ont été étudiées notamment dans [Lat4]. Nous allons considérer maintenant des généralisations n-aires, $n \geq 2$, des opérations chevron ou crochet ouvert. Les propriétés que nous établissons pour ces opérations, appelées respectivement chevron et crochet d'ordre n et notés Chn et COn sont identiques à celles démontrées dans le chapitre 3. Ces opérations sont en effet syntaxiques de degré 0 et le produit de n langages est indépendant des opérateurs associés.

L'intérêt primordial de cette étude réside dans la caractérisation des "equal matrix languages" d'ordre n définis par R. Siromoney [Si1] comme le cône rationnel engendré par le crochet ouvert d'ordre n des langages rationnels. Nous montrons donc que ces langages peuvent être caractérisés par des opérations unaires, alors que Greibach et Klingenstein [Gr3,Klin] n'en avaient donnés une caractérisation algébrique qu'avec l'opération binaire appelée "homomorphic duplication".

L'utilisation de l'opération crochet ouvert d'ordre n est surtout fructueuse parce qu'elle nous permet de conclure que les "equal matrix languages" d'ordre n, et donc la famille $\mathcal{L}'$ des "equal matrix languages", est effaçable, i.e. pour tout système générateur $\mathcal{S}$, on a

$$\mathcal{C}^f(\mathcal{S}) = \mathcal{C}(\mathcal{S}) = \mathcal{L}'.$$

Ceci constitue, à notre connaissance, le second exemple de hiérarchie de cônes rationnels de langages classique vérifiant cette propriété. Le premier exemple est en effet[GGG] la hiérarchie $\mathcal{F}_{m,n}$ des langages algébriques (m,n)-bornés définie dans [Gr1]. Ceci nous conduit à souligner deux faits. D'une part nos résultats sont établis de façon purement algébrique contrairement à [GGG], d'autre part les langages $K_n = \{(wc)^n w | w \in \{a,b\}^*\}$ générateurs de COn (Rat) sont des langages IRS [GGG], i.e. ne contiennent pas de langage rationnel infini. Ceci nous conduit à supposer que les propriétés de [GGG] sur les langages algébriques IRS doivent se généraliser à plusieurs autres familles de langages classiques.

B) PROPRIETES GENERALES

Définition 5.1

Etant donnés n langages $L_1,\ldots,L_n$ et X un alphabet d'au moins deux lettres disjoint de $X_{L_1} \cup \ldots \cup X_{L_n}$, nous désignerons par :

- Crochet ouvert d'ordre n de $L_1,\ldots,L_n$ le langage

$$COn(L_1,\ldots,L_n) = \{w\ell_1 w\ell_2 \ldots w\ell_n w \mid \ell_i \in L_i,\ 1 \leq i \leq n,\ w \in X^*\}.$$

- Chevron d'ordre n de $L_1,\ldots,L_n$ le langage

$$Chn(L_1,\ldots,L_n) = \{a^p\ell_1 a^p\ell_2 \ldots a^p\ell_n a^p \mid \ell_i \in L_i,\ 1 \leq i \leq n,\ p \geq 0,\ a \in X\}.$$

Etant donné n familles de langages $\mathcal{L}_1,\ldots,\mathcal{L}_n$ nous désignerons par $COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ (resp. $Chn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$) les cônes rationnels :

$$COn(\mathcal{L}_1,\ldots,\mathcal{L}_n) = \mathcal{C}^U(\{COn(L_1,\ldots,L_n) \mid L_i \in \mathcal{L}_i,\ 1 \leq i \leq n\})$$

$$Chn(\mathcal{L}_1,\ldots,\mathcal{L}_n) = \mathcal{C}^U(\{Chn(L_1,\ldots,L_n) \mid L_i \in \mathcal{L}_i,\ 1 \leq i \leq n\}).$$

Les opérateurs COn et Chn sont donc associés aux opérateurs portant le même nom. Les opérateurs fidèles associés seront respectivement définis par

$$COn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n) = \mathcal{C}^{fU}(\{COn(L_1,\ldots,L_n) \mid L_i \in \mathcal{L}_i, 1 \leq i \leq n\})$$

$$Chn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n) = \mathcal{C}^{fU}(\{Chn(L_1,\ldots,L_n) \mid L_i \in \mathcal{L}_i, 1 \leq i \leq n\}).$$

On a immédiatement :

Lemme 5.1

Les opérations Chn et COn sont régulières pour tout $n \geq 1$.

Il en résulte donc

Corollaire 5.1

Si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels (fidèles) principaux de générateurs respectifs $L_1,\ldots,L_n$ alors $COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ et $Chn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ (resp. $COn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ et $Chj^f(\mathcal{L}_1,\ldots,\mathcal{L}_n)$) sont des cônes rationnels (fidèles) principaux de générateurs respectifs $COn(L_1,\ldots,L_n)$ et $Chn(L_1,\ldots,L_n)$.

Nous pouvons donner une caractérisation du cône $COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ analogue à celle que nous avons établie pour les opérateurs C_i et CO_i. Pour cela introduisons :

Définition 5.2

Etant donné n langages $L_1,\ldots,L_n$ et A un ensemble appartenant à Rat_{n+1}, nous désignerons par crochet ouvert d'ordre n de $L_1,\ldots,L_n$ avec A le langage :

$$COn(L_1,\ldots,L_n|A) = \{a_1\ell_1a_2\ell_2\ldots a_n\ell_na_{n+1} | \ell_i \in L_i,\ 1 \leq i \leq n,$$

$$(a_1,\ldots,a_{n+1}) \in A\}.$$

Dans le cas des cônes rationnels nous avons alors le théorème :

Théorème 5.1

Soient $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnels (fidèles). La famille $COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ (resp. $COn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n)$) est identique à la ferme-

ture par union de la famille $\mathcal{L}'$ (resp. $\mathcal{L}''$) où

$$\mathcal{L}' = \{CON(L_1,\ldots,L_n|A) | L_i \in \mathcal{L}_i,\ 1 \leq i \leq n,\ A \in Rat_{n+1}\}$$

(resp. $\mathcal{L}'' = \{CON(L_1,\ldots,L_n|A) | L_i \in \mathcal{L}_i,\ 1 \leq i \leq n,$

$A = \{(\varphi_1 f,\ldots,\varphi_{n+1} f) | f \in K\},\ K \in Rat,\ \varphi_i$ homomorphisme

non effaçant, $1 \leq i \leq n+1\}$).

Démonstration

Comme dans le théorème 3.9 il nous suffit de vérifier que la fermeture par union de $\mathcal{L}'$ est un cône rationnel :

1) Soit φ un homomorphisme de monoïdes libres alors si $L' = CON(L_1,\ldots,L_n|A)$ appartient à $\mathcal{L}'$, alors $\varphi(L') = CON(\varphi(L_1),\ldots,\varphi(L_n)|\varphi(A))$ appartient à $\mathcal{L}'$ puisque $\varphi(L_i) \in \mathcal{L}_i$, $1 \leq i \leq n$ et $\varphi(A) \in Rat_{n+1}$.

2) Soit φ un homomorphisme alphabétique de U^* dans T^*, avec $L' = CON(L_1,\ldots,L_n|A) \subseteq T^*$. Alors $\varphi^{-1}(L') = CON(\varphi^{-1}(L_1),\ldots,\varphi^{-1}(L_n)|\varphi^{-1}(A))$ appartient à $\mathcal{L}'$ puisque $\varphi^{-1}(L_i) \in \mathcal{L}_i$, $1 \leq i \leq n$ et $\varphi^{-1}(A)$ appartient à Rat_{n+1} d'après la proposition 4.11.

3) Soit, comme dans la preuve du théorème 3.9, $K \in Rat(X^*)$ un langage régulier accepté par un automate fini auquel on associe les ensembles

$$K_m = \{w \in X^* | \delta(q_0,w) = q_m\}$$

$$K_{m,n} = \{w \in X^* | \delta(q_m,w) = q_n\}$$

$$K'_n = \{w \in X^* | \delta(q_n,w) \in F\}.$$

Soit $J = \{0,1,\ldots,n | q_0,\ldots,q_n \in Q\}$. Pour toute suite $\sigma = \{p_1,\ldots p_{2n}\}$

d'éléments de J soit :

$$A_\sigma = K_{p_1} \times K_{p_2,p_3} \times K_{p_3,p_4} \times \ldots \times K_{p_{2n-2},p_{2n-1}} \times K'_{p_{2n}}$$

Si $L' = COn(L_1,\ldots,L_n|A) \in \mathcal{L}'$, $L' \subseteq X^*$ alors

$$L' \cap K = \bigcup_\sigma COn(L_1 \cap K_{p_1,p_2},\ldots,L_n \cap K_{p_{2n-1},p_{2n}}|A \cap A_\sigma)$$

Or pour tout i, $K_{p_{2i-1},p_{2i}}$ est rationnel donc $L_1 \cap K_{p_{2i-1},p_{2i}} \in \mathcal{L}_i$, $1 \leq i \leq n$. De plus si $A = \{(\varphi_1 h,\ldots,\varphi_{n+1} h)|h \in R\} \in Rat_{n+1}$ alors

$$A \cap A_\sigma = \{(\varphi_1 h,\ldots,\varphi_{n+1} h)|h \in R \cap \varphi_1^{-1} K_{p_1} \cap \varphi_2^{-1} K_{p_1,p_2} \cap \ldots$$

$$\cap \varphi_n^{-1} K_{p_{2n-2},p_{2n-1}} \cap \varphi_{n+1}^{-1} K'_{p_{2n}}\}$$

et donc $A \cap A_\sigma \in Rat_{n+1}$.

Chacun des crochets ouverts d'ordre n appartient donc à $\mathcal{L}'$, et $L' \cap K$, leur union finie sur σ, appartient à la fermeture par union de $\mathcal{L}'$.

Donc si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels (fidèles) la fermeture par union de $\mathcal{L}'(\mathcal{L}")$ est un cône rationnel (fidèle). Pour achever la preuve il nous suffit de montrer que tout élément de $\mathcal{L}'(\mathcal{L}")$ appartient à $COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ (resp. $COn^f(\mathcal{L}_1,\ldots,\mathcal{L}_n)$). La preuve est identique au début de la démonstration du lemme 3.2 aussi ne la recommencerons-nous pas.

∎

Remarque

Par le même procédé on démontre que si $\mathcal{L}_1,\ldots,\mathcal{L}_n$ sont des cônes rationnels fidèles fermés par homomorphisme à effacement linéaire alors $COn^{lin}(\mathcal{L}_1,\ldots,\mathcal{L}_n)$, le plus petit cône rationnel fidèle fermé par homomorphisme à effacement linéaire contenant la famille $\{COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)|L_i \in \mathcal{L}_i,\ 1 \leq i \leq n\}$ est égal à la fermeture par union de la famille $\mathcal{L}'$. Contrairement à ce qui se passe pour l'opération crochet ouvert, le crochet ouvert d'ordre n, $n \geq 2$, ne semble pas préserver les langages effaçables. Nous pouvons seulement énoncer :

Définition 5.2 bis

Un langage L est linéairement effaçable si $\mathcal{E}(L) = \mathcal{E}^{lin}(L)$, où $\mathcal{E}^{lin}(L)$ désigne le plus petit cône rationnel fidèle contenant L et fermé par homomorphisme à effacement linéaire.

On a alors :

Corollaire 5.2

Pour tous langages linéairement effaçables $L_1,\ldots,L_n$, le langage $COn(L_1,\ldots,L_n)$ est linéairement effaçable.

Proposition 5.1

Pour tout $n \geq 1$, les opérations Chn et COn sont indépendantes de Rat (Rat_0).

Démonstration

Elle est immédiate puisque la condition $Chn(L_1,\ldots,L_n) \in Rat\,\mathcal{L}$ implique trivialement $Ch(L_1) \in Rat\,\mathcal{L}$. Le résultat découle alors du lemme 1.2.

∎

Afin d'établir que Chn et COn sont des opérations syntaxiques de degré 0, montrons tout d'abord

Proposition 5.2

Soient $L \subseteq X^*$ et $L' \subseteq X'^*$ deux langages strictement disjoints et $\mathcal{L}_1,\ldots,\mathcal{L}_n$, $n \geq 1$, des cônes rationnels fermés par union. Alors la condition $L.L' \in CO(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ implique $L \in \mathcal{L}_1$ ou $L' \in CO(n-1)(\mathcal{L}_2,\ldots,\mathcal{L}_n)$.

Démonstration

Soient L et L' tels que $LL' \in COn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$. D'après le théorème 5.1 il existe donc $p \in \mathbb{N}$, pn langages $L_{ij} \in \mathcal{L}_i$, $1 \leq i \leq n$, $1 \leq j \leq p$ et p ensembles $A_j \in Rat_{n+1}$, $1 \leq j \leq p$ tels que :

$$LL' = \bigcup_{j=1}^{p} COn(L_{1j},\ldots,L_{nj} \mid A_j)$$

Nous supposerons que pour tout j, les langages $L_{1j},\ldots,L_{nj}$ et les ensembles A_j sont non vides.

Pour tout $x \in L$ et $y \in L'$ il existe donc j, $1 \leq j \leq p$, $(a_1,\ldots,a_{n+1}) \in A_j$ et $w_i \in L_{ij}$, $1 \leq i \leq n$ tels que

$$xy = a_1w_1a_2w_2\ldots a_nw_na_{n+1}$$

De même que dans la preuve du théorème 3.10 deux cas sont possibles :

a) $\forall x \in L$, $\exists y \in L'$, $1 \leq j \leq p$, $(a_1,\ldots,a_{n+1}) \in A_j$ et $w_i \in L_{ij}$, $1 \leq i \leq n$ tels que $xy = a_1w_1\ldots a_nw_na_{n+1}$ et x est un facteur gauche propre de a_1w_1. Montrons que dans ce cas

$$L = L_0 = \{a_1f \in X^* \mid \exists z \in X'X'^*,\ a_2,\ldots,a_{n+1} \in X'^*,\ 1 \leq j \leq n :$$

$$(a_1,a_2,\ldots,a_{n+1}) \in A_j,\ fz \in L_{1j}\} \cup \{a_1' \in X^* \mid \exists a_1'' \in X'X'^* \mid a_1'a_1'' \in \pi_1(A_j)\}.$$

Les conditions de ce premier cas impliquent $L \subseteq L_0$. Réciproquement si $w \in L_0$ alors soit il existe $a_1'' \in X'^*X'$ tel que $wa_1'' \in \pi_1(A_j)$ et donc il existe $a_2,\ldots,a_{n+1}$ et des mots $w_i \in L_{ij}$, $1 \leq i \leq n$, puisque ces langages sont non vides tels que $wa_1''w_1a_2\ldots w_na_{n+1} \in COn(L_{1j},\ldots,L_{2j} \mid A_j) \subseteq L.L'$ et la disjonction des alphabets implique $w \in L$, soit $w = a_1f$ et l'on aboutit à la même conclusion de façon analogue. Donc $L = L_0$.

Or si $\quad P_{1j} = \{w \in X^* \mid \exists z \in X'X'^*, wz \in L_{1j}\}$

et $\quad A'_{1j} = \{w \in X^* \mid \exists z \in X'X'^*, wz \in \pi_1(A_j)\}$

$P_{1j} \in \mathcal{L}_1$ et $A'_{1j} \in$ Rat puisque ces langages sont images de L_{1j} et $\pi_1(A_j)$ dans la transduction rationnelle qui n'accepte que des mots appartenant à $X^*X'X'^*$ et efface le suffixe dans $X'X'^*$ à ceux-ci. Comme on a de façon évidente

$$L = L_0 = \bigcup_{j=1}^{p} \pi_1(A_j)P_{1j} \cup \left(\bigcup_{j=1}^{p} A'_{1j}\right)$$

et que le cône rationnel $\mathcal{L}_1$ est fermé par union, il en résulte que $L \in \mathcal{L}_1$.

b) $\exists x \in L$ tel que pour tout $y \in L'$, $1 \le j \le p$, $(a_1,\ldots,a_{n+1}) \in A_j$, $w_i \in L_{ij}$ vérifiant $xy = a_1w_1a_2w_2\ldots w_na_{n+1}$, alors a_1w_1 est facteur gauche de x. Montrons alors que L' est égal au langage

$$L'_0 = \{zw_ka_{k+1}\ldots w_na_{n+1} \in X'^* \mid 2 \le k \le n, \exists a_1,\ldots,a_{k-1},f, w_1,\ldots,w_{k-1} \in X^*,$$

$$1 \le j \le p : a_1w_1a_2\ldots w_{k-1}f = x, (a_1,\ldots,a_{k-1},fz,a_{k+1},\ldots,a_{n+1}) \in A_j$$

$$w_i \in L_{i,j}, 1 \le i \le n\}.$$

Les conditions de ce cas impliquent $L' \subseteq L'_0$. Réciproquement si $w \in L'_0$, alors il existe $a_1,\ldots,a_{k-1}$, $f \in X^*$, $w_i \in L_{i,j}$, $1 \le i \le k-1$, pour un certain j, $1 \le j \le p$ et k, $2 \le k \le n$ tels que

$$a_1w_1\ldots a_{k-1}w_{k-1}fw \in \mathrm{COn}(L_{1j},\ldots,L_{nj}|A_j) \subseteq LL'$$

et la disjonction des alphabets implique $w \in L'$ i.e. $L'_0 \subseteq L'$.

Pour tout suffixe w de x et tout i et j, $1 \le i \le n$, $1 \le j \le p$, soient alors les ensembles :

$$A_{ij}(w) = \{(z,a_{i+1},\ldots,a_{n+1}) \in (X'^*)^{n-i+1} \mid \exists a_1,\ldots,a_{i-1},f \in X^*,$$

$$w_k \in L_{k,j},\ 1 \leq k \leq i-1 : a_1 w_1 \ldots w_{i-1} f = w,$$

$$(a_1,\ldots,a_{i-1},fz,a_{i+1},\ldots,a_{n+1}) \in A_j\}.$$

L'ensemble des $w_1,\ldots,w_{k-1}$ tels qu'il existe $a_1,\ldots,a_{i-1},f$ vérifiant $a_1 w_1 \ldots w_{k-1} f = w$ est fini. Il en résulte alors, par un procédé analogue à celui utilisé pour démontrer la proposition 4.10 et après application de la proposition 4.9, que $A_{ij}(w) \in Rat_{n-i+1}$.

Donc pour tout i, j et w facteur droit de x, l'ensemble $CO(n-i)(L_{i+1},\ldots,L_n|A_{ij}(w))$ est dans $CO(n-i)(\mathcal{L}_{i+1},\ldots,\mathcal{L}_n)$, lui même inclus dans $CO(n-1)(\mathcal{L}_2,\ldots,\mathcal{L}_n)$.

Or nous avons

$$L' = L'_0 = \cup\ (CO(n-i)(L_{i+1},\ldots,L_n\ |A_{ij}(w))|w \text{ facteur droit de } x,\ 1 \leq i < n,\ 1 \leq j \leq p\}.$$

Donc $L' \in CO(n-1)(\mathcal{L}_2,\ldots,\mathcal{L}_n)$ puisque ce cône rationnel est fermé par union.

∎

Il en découle immédiatement :

Corollaire 5.3

Soient $L \subseteq X^*$ et $L' \subseteq X'^*$ deux langages et $\mathcal{L}_1,\ldots,\mathcal{L}_n$ des cônes rationnels fermés par union tels que $L.L' \in Chn(\mathcal{L}_1,\ldots,\mathcal{L}_n)$ alors $L \in \mathcal{L}_1$ ou $L' \in CO(n-1)(\mathcal{L}_2,\ldots,\mathcal{L}_n)$.

On peut alors énoncer

Théorème 5.2

Pour tout $n \geq 1$, les opérations Chn et COn sont syntaxiques de degré 0.

Démonstration

D'après la proposition 5.1 il nous suffit de montrer que Chn et COn sont de degré 0. Faisons la preuve pour COn, elle est identique pour Chn. Soient donc $L_1,\ldots,L_n$ des langages strictement disjoints et $\mathscr{L}_1,\ldots,\mathscr{L}_n$ des cônes rationnels (fidèles) tels que $COn(L_1,\ldots,L_n) \in COn(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ $(COn^f(\mathscr{L}_1,\ldots,\mathscr{L}_n))$. La disjonction des alphabets implique trivialement que $L_1\ldots L_n$ appartient à $COn(\mathscr{L}_1,\ldots,\mathscr{L}_n)$ $(COn^f(\mathscr{L}_1,\ldots,\mathscr{L}_n))$. Par application répétée de la proposition précédente il en résulte que soit il existe i, $1 \leq i \leq n-1$, tel que $L_i \in Rat$ soit $L_n \in \mathscr{L}_n$. Donc COn est de degré 0.

∎

Des propositions 5.1, 5.2 et du théorème précédent il résulte donc :

Théorème 5.3

Soient $\mathscr{L}_1,\ldots,\mathscr{L}_n$ des cônes rationnels (fidèles) fermés par union tels que $\mathscr{L}_1 \vee \ldots \vee \mathscr{L}_n$ n'est pas fermé par crochet ouvert d'ordre n (resp. COn^f, Chn, Chn^f). Alors :

1) Le plus petit cône rationnel (fidèle) contenant $\mathscr{L}_1 \vee \ldots \vee \mathscr{L}_n$ et fermé par COn (resp. Chn, COn^f, Chn^f) est non principal.

2) Le plus petit cône rationnel (fidèle) contenant $\mathscr{L}_1 \vee \ldots \vee \mathscr{L}_n$ et fermé par produit et COn (resp. Chn, COn^f, Chn^f) est non principal.

3) La plus petite FAL (fidèle) contenant $\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n$ fermée par COn (resp. Chn, COn^f, Chn^f) est non principale.

4) Pour $t \in \{\mathrm{COn}, \mathrm{Chn}\}$ (resp. COn^f, Chn^f) on a :

$$\mathcal{E}_t^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) \subsetneq \underline{\mathrm{Rat}}\ \mathcal{E}_t^U(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) \subsetneq \mathcal{F}_t(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n)$$

$$(\text{resp. } \mathcal{E}_t^{fU}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) \subsetneq \underline{\mathrm{Rat}}_0\ \mathcal{E}_t^{fU}(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n) \subsetneq \mathcal{F}_t^f(\mathcal{L}_1 \vee \ldots \vee \mathcal{L}_n))$$

C) APPLICATIONS

Les opérateurs COn, $n \geq 1$, nous permettent de donner une nouvelle définition des "equal matrix languages" introduits par R. Siromoney [Si1].

S. Greibach [Gr3] et Klingenstein [Klin] ont caractérisé ces langages à l'aide des opérations binaires "homomorphic duplication", nous montrons ainsi que des opérations unaires suffisent, et que les "equal matrix languages" d'ordre n forment des cônes rationnels linéairement effaçables.

Définition 5.3

Soit $X = \{a,b\}$ un alphabet et c un nouveau symbole.

Pour tout $n \geq 1$, soient les langages

$$K_n = \{(wc)^n w \mid w \in X^*\}$$

$$L_n = \{(a^p c)^n a^p \mid p \geq 0\}$$

Du corollaire 5.1 on tire immédiatement

Proposition 5.3

Pour tout $n \geq 1$, K_n et L_n sont des générateurs respectifs des cônes rationnels (fidèles) COn(Rat) et Chn(Rat) (resp. $COn^f(Rat)$ et $Chn^f(Rat)$).

Comme le langage rationnel {c} est effaçable il résulte du corollaire 5.2 :

Proposition 5.4

Pour tout $n \geq 1$, les langages K_n sont linéairement effaçables.

Enfin des propositions 5.1, 5.2 et du théorème 5.3 on tire :

Proposition 5.5

Pour tout $n \geq 1$, le plus petit cône rationnel fermé par COn (resp. Chn) est non principal et n'est pas fermé par produit. La FAL engendrée par ce cône n'est pas fermée par COn (resp. Chn), n'est pas principale et est indécomposable. De plus $\mathcal{C}^{f}_{COn}(Rat)$ est égal à $\mathcal{C}_{COn}(Rat)$.

Etant donné la caractérisation de [Klin] il résulte que $COn(Rat) = \mathcal{C}(Kn)$ est exactement la famille des "equal matrix languages" d'ordre n [Sil]. On retrouve donc que les "equal matric languages" d'ordre n forment un cône rationnel principal tandis que $\mathcal{L}_{EM} = \bigcup_{n \geq 1} COn(Rat)$ est un cône rationnel non principal.

La proposition 5.5 nous permet enfin de conclure que les "equal matrix languages" d'ordre n forment des cônes rationnels linéairement effaçables et donc $\mathcal{L}'$ la famille des "equal matrix languages" est aussi linéairement effaçable. A notre connaissance ce résultat n'avait jamais été établi.

Chapitre 6

Indépendance des Opérations

INTRODUCTION

Le nombre croissant d'opérations sur les langages qui sont utilisés pour étudier les langages formels conduit naturellement à se poser le problème de leur indépendance. De nombreux articles sont consacrés à cette étude [GH,Gr6,Ito,BG,...] qui permet notamment de comparer certaines familles de langages classiques.

Dans ce chapitre nous comparons entre elles les opérations chevron, crochet et crochet ouvert puis aux opérations étoile propre, "homomorphic replication" [Gr3] et substitution.

Après avoir donné une définition de l'indépendance des opérations, nous déduisons des chapitres précédents un certain nombre de résultats généraux puis nous montrons que le chevron, le crochet et le crochet ouvert sont indépendants de l'étoile propre et réciproquement. Nous établissons ensuite que "l'homomorphic replication" et la substitution sont indépendants de ces opérations tandis que le chevron, crochet et crochet ouvert ne sont indépendants que des opérations substitution dans des cônes rationnels particuliers appelés 1-itératifs.

Les résultats établis ici seront utilisés dans le chapitre suivant pour comparer diverses hiérarchies de langages vérifiables.

Définition 6.1

Soient t et t' deux opérations sur les langages.

L'opération t' est <u>indépendante</u> de t si pour tout cône rationnel $\mathcal{L}$ la condition $\mathcal{C}^{U}_{t'}(\mathcal{L}) \subseteq \mathcal{C}^{U}_{t}(\mathcal{L})$ implique $\mathcal{C}^{U}_{t'}(\mathcal{L}) \subseteq \mathcal{L}$.

Du corollaire 2.12 sur les opérations indépendantes de <u>Rat</u> il résulte alors :

Corollaire 6.1

Toute opération t indépendante de Rat est indépendante de l'étoile propre.

Démonstration

Pour tout cône rationnel $\mathcal{L}$ si $\mathcal{C}_t^U(\mathcal{L}) \subseteq \underline{\text{Rat}}\ \mathcal{L}$ alors pour tout langage $L \in \mathcal{C}_t^U(\mathcal{L})$, $t(L) \in \mathcal{C}_t^U(\mathcal{L}) \subseteq \underline{\text{Rat}}\ \mathcal{L}$ et donc $L \in \mathcal{L}$ puisque t est indépendant de Rat.

■

De même du lemme 2.8 on tire

Corollaire 6.2

Soit t une opération associée à l'opérateur T telle que le produit de deux langages est indépendant de T. Alors l'opération étoile (propre) est indépendante de t.

Démonstration

Si $(Lc)^+$ appartient à $\mathcal{C}_t^U(\mathcal{L})$ alors il existe n tel que $(Lc)^+ \in T^n(\mathcal{L})$ soit donc $((Lc)^+))^{n+1} = (Lc)^+ \in T^n(\mathcal{L})$ et d'après le lemme 2.8 $(Lc)^+ \in \mathcal{L}$.

■

En fait ce corollaire montre que si le produit de deux langages est indépendant de T alors toute FAL incluse dans $\mathcal{C}_T^U(\mathcal{L})$ est incluse dans $\mathcal{L}$.

En particulier nous en déduisons, par exemple, que pour tout cône

rationnel $\mathcal{L}$, $\mathcal{E}^{U}_{\odot}(\mathcal{L})$ forme une FAL si et seulement si $\mathcal{L}$ est FAL. Ce résultat est notablement plus fort que celui de [GH] ou [Gr3] où les auteurs établissent seulement qu'il existe des cônes rationnels fermés par produit qui ne sont pas fermés par étoile.

Avant d'étudier de nouvelles opérations, énonçons tout d'abord le résultat suivant qui résulte directement de l'étude du crochet, du chevron et du crochet ouvert faite précédemment ainsi que de [Lat4] :

Proposition 6.1

1) L'opération crochet est fortement indépendante des opérations chevron et crochet ouvert.

2) Le crochet ouvert est fortement indépendant du chevron et du crochet.

3) L'opération étoile propre est fortement indépendante du chevron, du crochet et du crochet ouvert.

Notons que les relations d'indépendance ne sont pas symétriques : en effet le chevron n'est pas indépendant du crochet, ni du crochet ouvert.

Rappelons maintenant la définition des opérations "homomorphic replication" introduites dans [GS3] comme des opérations unaires et étudiées dans [Gr3] sous une forme binaire qui se révèle être une opération binaire syntaxique au sens de [Gr3]. En fait on démontre sans difficulté que l'on peut considérer une opération unaire syntaxique au sens du chapitre 2 qui engendre les mêmes familles de langages.

Définition 6.2

Soit p une application de $\{1,2,\ldots,n\}$ dans $\{1,R\}$ et $h_1,\ldots,h_n$ des homomorphismes. Soit $w^R = \tilde{w}$ pour tout mot w. Soit $f_{p,h_1,\ldots,h_n}$ l'opération unaire définie par

$$f_{\rho,h_1,\ldots,h_n}(L) = \{f_1(w)^{\rho(1)}\ldots h_n(w)^{\rho(n)} | w \in L\}$$

et soit, pour toute famille de langages $\mathcal{L}$, ρ l'opérateur défini par

$$\rho(\mathcal{L}) = \{f_{\rho,h_1,\ldots,h_n}(L) | L \in \mathcal{L}, h_1,\ldots,h_n \text{ des homomorphismes}\}.$$

L'opérateur ρ est appelé opérateur "d'homomorphic replication" d'ordre n [GS3,Gr3].

On montre alors sans peine que

Proposition 6.2

Pour tout $n \geq 2$ et pour toute application ρ de $\{1,\ldots,n\}$ dans $\{1,R\}$ l'opérateur ρ est associé à l'opération syntaxique unaire ρ de degré 0 définie par

$$(L) = \{w^{\rho(1)}\ldots w^{\rho(n)} | w \in L\}.$$

De plus le produit de deux langages est indépendant de ρ.

Par abus de langage l'opération ρ sera également appelée "homomorphic replication" d'ordre n. Nous dirons que $f_{\rho,h_1,\ldots,h_n}$ est <u>non effaçant</u> si les homomorphismes $h_1,\ldots,h_n$ sont non effaçants. $f_{\rho,h_1,\ldots,h_n}$ sera à <u>effacement linéaire</u> si pour tout hi, $1 \leq i \leq n$, il existe $k > 0$ tel que pour tout w, $|w| \geq k$, on a $|w| \leq k|h(w)|$.

Comparons les opérations Ch, C et CO aux opérations "homomorphic replication". On a tout d'abord

Lemme 6.1

Soit $\mathcal{L}$ un cône rationnel fidèle fermé par union et par "homomorphic replication". Alors $\mathcal{L}$ est fermé par chevron, crochet et crochet ouvert.

Démonstration

Soit $L \subseteq X^*$ un langage quelconque de $\mathcal{L}$ et a et b deux symboles n'appartenant pas à X. Trivialement le langage $\{a,b\}^* L$ appartient à $\mathcal{L}$. Soient alors h_1 et h_2 les homomorphismes définis par $h_1(x) = x$, $x \in \{a,b\}$ et $h_1(x) = \varepsilon$ si $x \in X$, $h_2(x) = x$ si $x \in X$ et $h_2(x) = \varepsilon$ si $x \in \{a,b\}$. On a alors de façon immédiate :

$$C(L) = \{h_1(w)h_2(w)h_1(w)^R \mid w \in \{a,b\}^* L\} \text{ et}$$

$$CO(L) = \{h_1(w)h_2(w)h_1(w) \mid w \in \{a,b\}^* L\}, \text{ i.e.}$$

$\mathcal{L}$ est fermé par crochet et crochet ouvert -donc par chevron- si $\mathcal{L}$ est fermé par "homomorphic replication".

■

Notons que le lemme précédent reste vrai si l'on suppose seulement que $\mathcal{L}$ est fermé par "homomorphic replication" à effacement linéaire.

Inversement nous allons montrer que l'opération "homomorphic replication" est fortement indépendante du crochet et du crochet ouvert et donc du chevron. Pour cela procédons de manière analogue à [Gr6] :

Définition 6.3

Soit $L \subseteq X^*$ un langage. On dira que L vérifie la propriété 1 si pour tout $xyz \in L$ soit $\text{card}(L \cap X^* y X^*) \leq 1$ soit $\text{card}(L \cap xX^* z) \leq 1$. On a alors

Lemme 6.2

Soit $\mathcal{L}$ une famille de langages contenant les parties finies, fermée par union et par produit avec les ensembles unitaires.

Si L est un langage ayant la propriété 1 tel que

$$L = \bigcup_{j=1}^{n} C(M_j,A_j) \text{ avec } A_j \in Rat_2 \text{ et } M_j \in \mathcal{L},$$

$1 \leq j \leq n$ (resp. $L = \bigcup_{j=1}^{n} CO(M_j,A_j)$), alors $L \in \mathcal{L}$.

Démonstration

Soit $w \in L$. Alors $\exists j \in [1,n]$, $(a_j,b_j) \in A_j$ et $m_j \in M_j$ tel que $w = a_j m_j \tilde{b}_j$. Considérons alors $x = a_j$, $y = m_j$ et $z = \tilde{b}_j$. Alors

1) Soit $card(L \cap a_j X^* \tilde{b}_j) \leq 1$ i.e. $card(M_j) = 1$.

2) Soit $card(L \cap X^* m_j X^*) \leq 1$ i.e. $card(A_j) = 1$.

- Si $card(A_j) = 1$, i.e. $A_j = \{(a,b)\}$, alors $\{a\}\ M_j\ \{b\} \in \mathcal{L}$.

- Si $card(M_j) = 1$ démontrons que A_j est fini. En effet sinon il existe $(\alpha,\beta) \in A_j$ avec $|\alpha\beta|$ assez grand et tel que $\alpha = \alpha_1\alpha_2\alpha_3$, $\beta = \beta_1\beta_2\beta_3$ avec $|\alpha_2\beta_2| > 0$ et vérifiant $(\alpha_1\alpha_2^k\alpha_3,\ \beta_1\beta_2^k\beta_3) \in A_j$. Ainsi si $M_j = \{m\}$ nous aurons $\forall k \geq 0$; $w_k = \alpha_1\alpha_2^k\alpha_3 m \tilde{\beta}_3 \tilde{\beta}_2^k \tilde{\beta}_1 \in L$. Si l'on pose alors $\alpha_1\alpha_2 = x$

$$\alpha_2^{k-1}\alpha_2\ \tilde{m}\ \beta_3\ \tilde{\beta}_2^{k-1} = y$$

$$\tilde{\beta}_2\ \tilde{\beta}_1 = z$$

alors $L \cap X^*yX^* = \{\alpha_1\alpha_2^m\alpha_3 m \tilde{\beta}_3 \tilde{\beta}_2^m \tilde{\beta}_1 \mid n \in \mathbb{N}\} = L \cap xX^*y$, ce qui contredit notre hypothèse. Donc A_j est fini et par conséquent $\{\alpha m \tilde{\beta} \mid (\alpha,\beta) \in A_j\} \in \mathcal{L}$.

La preuve est analogue pour le crochet ouvert.

∎

Lemme 6.3

Soit $\mathcal{L}$ un cône rationnel fidèle, $L \subseteq X^*$ un langage dans $\mathcal{L}$ et c un nouveau symbole. Pour tout $k \geq 3$, $t_1,\ldots,t_k$ dans $\{1,R\}$, le langage $L_k = \{w^{t_1}cw^{t_2}c\ldots cw^{t_k} \mid w \in L\}$ est dans $C^f(\mathcal{L})$ (resp. $\mathcal{E}^f_{CO}(\mathcal{L})$) si et seulement si $L_k \in \mathcal{L}$. Si $\mathcal{L}$ est un cône rationnel alors $L_k \in C(\mathcal{L})$ (resp. $CO(\mathcal{L})$ si et seulement si $L_k \in \mathcal{L}$.

Démonstration

L_k vérifie la propriété 1. Si $L_k \in C^f(\mathcal{L})$ alors le résultat découle du lemme 6.2 et de la proposition 1.1. Si $L_k \in C^f_{CO}(\mathcal{L})$ alors il existe $p \in \mathbf{N}$ tel que $L_k \in CO^p(\mathcal{L})$ et l'application répétée du lemme précédent et du théorème 3.1 conduit au résultat. Il en est de même pour les opérateurs C et CO.

■

On peut alors énoncer

Théorème 6.1

Soient $\mathcal{L}_1$ et $\mathcal{L}_2$ deux cônes rationnels fidèles fermés par union, avec $\mathcal{L}_1$ fermé par homomorphisme à effacement linéaire.

1) $\mathcal{E}^f_\rho(\mathcal{L}_2) \subseteq C^f(\mathcal{L}_1)$ (resp. $\mathcal{E}^f_\rho(\mathcal{L}_2) \subseteq \mathcal{E}^f_{CO}(\mathcal{L}_1)$) si et seulement si $\mathcal{E}^f_\rho(\mathcal{L}_2) \subseteq \mathcal{L}_1$,

et $\mathcal{E}_\rho(\mathcal{L}_2) \subseteq C(\mathcal{L}_1)$ (resp. $\mathcal{E}_\rho(\mathcal{L}_2) \subseteq \mathcal{E}_{CO}(\mathcal{L}_1)$) si et seulement si $\mathcal{E}_\rho(\mathcal{L}_2) \subseteq \mathcal{L}_1$.

2) Si $\mathcal{L}_1$ n'est pas fermé par "homomorphic replication" non effaçant, alors

$$\mathcal{E}^f_\rho(\mathcal{L}_1) - C^f(\mathcal{L}_1) \neq \emptyset$$

$$(\text{resp. } \mathcal{E}^f_\rho(\mathcal{L}_1) - \mathcal{E}^f_{CO}(\mathcal{L}_1) \neq \emptyset),$$

et si $\mathcal{L}_1$ est un cône rationnel

$$\mathcal{E}_\rho(\mathcal{L}_1) - C(\mathcal{L}_1) \neq \emptyset$$

$$(\text{resp. } \mathcal{E}_\rho(\mathcal{L}_1) - \mathcal{E}_{CO}(\mathcal{L}_1) \neq \emptyset).$$

3) $C^f(\mathcal{L}_1)$ (resp. $\mathcal{E}^f_{CO}(\mathcal{L}_1)$) est fermé par "homomorphic replication" non effaçant si et seulement si $\mathcal{L}_1 = \mathcal{E}^f_\rho(\mathcal{L}_1) = C(\mathcal{L}_1)$ (resp. $\mathcal{L}_1 = \mathcal{E}^f_\rho(\mathcal{L}_1) = \mathcal{E}^f_{CO}(\mathcal{L}_1)$) et si $\mathcal{L}_1$ est un cône rationnel alors $C(\mathcal{L}_1)$ (resp. $\mathcal{E}_{CO}(\mathcal{L}_1)$) est fermé par "homomorphic replication" si et seulement si $\mathcal{L}_1 = C(\mathcal{L}_1) = \mathcal{E}_\rho(\mathcal{L}_1)$ (resp. $\mathcal{L}_1 = \mathcal{E}_{CO}(\mathcal{L}_1) = \mathcal{E}_\rho(\mathcal{L}_1)$).

Démonstration

1) Supposons que $\mathcal{L} = \mathcal{E}^f_\rho(\mathcal{L}_2) \subseteq C^f(\mathcal{L}_1)$. Soit alors L un langage de $\mathcal{L}$ et c un nouveau symbole. Comme $\mathcal{L}$ est fermé par "homomorphic replication" non effaçant, le langage $L_3 = \{wcwcw \mid w \in L\}$ appartient à $C^f(\mathcal{L}_1)$ et donc à $\mathcal{L}_1$ d'après le lemme 6.3. Comme $\mathcal{L}_1$ est fermé par homomorphisme à effacement linéaire il en résulte que $L \in \mathcal{L}_1$. La preuve est identique pour les opérateurs CO^f, C et CO.

2) Il suffit de poser $\mathcal{L}_2 = \mathcal{L}_1$ et d'utiliser le point 1) ainsi

que pour tout cône rationnel $\mathscr{L}$, $C^f(\mathscr{L}) = C(\mathscr{L})$ (cf chapitres 1 et 3).

3) Du point 2) il résulte que si $\mathscr{L}_1 \neq \mathscr{E}^f_\rho(\mathscr{L}_1)$ alors $C^f(\mathscr{L}_1)$ n'est pas fermé par "homomorphic replication". Or un cône rationnel fermé par union est fermé par "homomorphic replication" si et seulement si il est fermé par "homomorphic replication" non effaçant, donc si $\mathscr{L}_1$ est un cône rationnel fermé par union tel que $\mathscr{L}_1 \neq \mathscr{E}_\rho(\mathscr{L}_1)$ alors $C(\mathscr{L}_1)$ n'est pas fermé par "homomorphic replication".

Si $\mathscr{L}_1 \neq C^f(\mathscr{L}_1)$ alors comme $C^f(\mathscr{L}_1) \subseteq \mathscr{E}^f_\rho(\mathscr{L}_1)$ on a $\mathscr{E}^f_\rho(\mathscr{L}_1)$ non inclus dans $\mathscr{L}_1$ et donc $\mathscr{E}^f_\rho(\mathscr{L}_1)$ n'est pas inclus dans $C^f(\mathscr{L}_1)$ d'après le point 1.

La preuve est identique pour CO^f, C et CO.

∎

Corollaire 6.3

Les opérations "homomorphic replication" sont indépendantes des opérations chevron, crochet et crochet ouvert.

Nous achèverons ce chapitre en comparant maintenant le chevron, le crochet et le crochet ouvert à la substitution.

On a tout d'abord :

Théorème 6.2

Soit $\mathscr{L}_1$ un cône rationnel et $\mathscr{L}_2$ un cône rationnel fidèle. Alors $C(\mathscr{L}_1)$ (resp. $C^f(\mathscr{L}_1)$, $\mathscr{E}^f_{CO}(\mathscr{L}_1)$, $\mathscr{E}_{CO}(\mathscr{L}_1)$) est fermé par $\underline{\mathscr{L}_2}$, l'opérateur de substitution dans $\mathscr{L}_2$, si et seulement si $\mathscr{L}_1 = C(\mathscr{L}_1) = \mathscr{R}(\mathscr{L}_1)$ (resp. $\mathscr{L}_1 = C^f(\mathscr{L}_1)$, $\mathscr{E}^f_{CO}(\mathscr{L}_1)$,

$\mathcal{C}_{CO}(\mathcal{L}_1))$ et $\mathcal{L}_1$ est fermé par $\underline{\mathcal{L}_2}$.

Démonstration

On sait [GG1] que $\underline{\mathcal{L}_2}$ Rat = $\mathcal{L}_2$ et que la substitution est associative. Donc

$$\underline{\mathcal{L}_2}\ C(\mathcal{L}_1) = (\underline{\mathcal{L}_2}\ \text{Rat})\ C(\mathcal{L}_1)$$

$$= \underline{\mathcal{L}_2}\ (\underline{\text{Rat}}\ C(\mathcal{L}_1))$$

Donc si $C(\mathcal{L}_1)$ est fermé par $\underline{\mathcal{L}_2}$, alors $C(\mathcal{L}_2)$ est une FAL. Or d'après [BCN] si $\mathcal{L}_1$ n'est pas fermé par crochet alors $C(\mathcal{L}_1)$ n'est pas une FAL. D'où le résultat. La preuve est semblable dans les autres cas.

∎

Corollaire 6.4

Les opérations substitution sont indépendantes des opérations chevron, crochet et crochet ouvert.

Nous ne considèrerons la réciproque que dans un cas particulier dont la preuve est identique à celle donnée par [BoNi1] pour le lemme 1.2. Introduisons tout d'abord une nouvelle définition :

Définition 6.4

Un langage L est fortement k-itératif, $k \geq 1$, si pour tout t il existe des constantes p et q telles que pour tout mot $w \in L$ tel que p occurrences de lettres de w au moins sont considérées comme distinguées, alors w se factorise en $w = x_1y_1 \ldots x_ky_kx_{k+1}$ avec :

1) Il existe i tel que x_i, y_i et x_{i+1} contiennent chacun au moins t

positions distinguées et $x_i y_i x_{i+1}$ contient au plus q positions distinguées.

2) Pour tout $n \geq 0$, le mot $x_1 y_1^n \ldots x_k y_k^n x_{k+1}$ appartient à $\mathcal{L}$.

Une famille de langages $\mathcal{L}$ est fortement (k-) itérative si tout $L \in \mathcal{L}$ est fortement (k-) itératif.

Cette définition, introduite par Greibach dans [Gr6] généralise celle de [Ogd] et permet de retrouver des résultats de [GS3,Gr7,Gr8] similaires à ceux de [VdW]. Il est connu que Rat et Alg sont fortement itératifs. On a alors la proposition suivante dont les points 1 et 2 sont établis dans [Gr6] :

Proposition 6.3

Soient $\mathcal{L}_1$ et $\mathcal{L}_2$ des cônes rationnels strictement itératifs, alors :

1) $\mathcal{E}_\rho(\mathcal{L}_1)$ est fortement itératif.

2) $\mathcal{E}_{\underline{\mathcal{L}}_1}(\mathcal{L}_2)$ est fortement itératif.

3) $C(\mathcal{L}_1)$, $\mathcal{E}_{Ch}(\mathcal{L}_1)$ et $\mathcal{E}_{CO}(\mathcal{L}_1)$ sont strictement itératifs.

Démonstration

Nous ne ferons la preuve que pour le crochet : elle est identique pour le crochet ouvert. Enfin comme $\mathcal{E}_{Ch}(\mathcal{L}_1) \subseteq C(\mathcal{L}_1)$ ceci prouvera le point 3).

Soit alors L un langage de $C(\mathcal{L}_1)$, il existe donc $A \in Rat_2$ et $M \in \mathcal{L}$ tels que $L = C(M,A)$. Par hypothèse M est fortement k-itératif, démontrons que L est fortement max(2,k)-itératif. Pour tout $t \geq 1$, soient p et q les entiers satisfaisant la définition 6.4 pour M et soient p' et q' les entiers associés au langage rationnel R associé à A par le lemme d'itération 4.7

sur les relations rationnelles. Soit alors $p_1 = p+p'$ et $q_1 = \max(q,q')$. Considérons un mot $w \in L$ dont p_1 occurrences de lettres au moins sont distinguées. Par définition il existe $(x,y) \in A$ et $m \in M$ tels que $w = xm\tilde{y}$. Si au moins p occurrences de lettres distinguées se trouvent dans m alors w admet une factorisation en $w = xu_1v_1 \ldots u_kv_ku_{k+1}\tilde{y}$ qui satisfait la définition 6.4. Sinon au moins p' occurrences de lettres distinguées sont dans $x\tilde{y}$ et, d'après le lemme 4.7 w admet la factorisation

$$w = x_1x_2x_3m\tilde{y}_3\tilde{y}_2\tilde{y}_1$$

telle que $x_1x_2^sx_3m\tilde{y}_3\tilde{y}_2^s\tilde{y}_1 \in L$ et satisfait la définition 6.4. D'où le résultat.

∎

Remarquons que l'on peut établir des résultats analogues pour les opérations crochet, chevron et crochet ouvert de degré i ainsi que pour le crochet ouvert et le chevron d'ordre n.

On peut alors énoncer la proposition suivante dont la preuve est identique en tous points à celle de [BoNi1] pour le lemme 1.2.

Proposition 6.4

Soit L un langage, $\mathcal{L}_1$ un cône rationnel fortement 1-itératif et $\mathcal{L}_2$ un cône rationnel (fidèle) fermé par union.
La condition $Ch(L) \in \underline{\mathcal{L}_1}\ \mathcal{L}_2$ implique que $L \in \mathcal{L}_2$.

D'où

Théorème 6.3

Soient $\mathcal{L}_1$ et $\mathcal{L}_2$ des cônes rationnels (fidèles) et $\mathcal{L}$ un cône rationnel fortement 1-itératif. Pour t = C, Ch et CO on a :

1) $\mathcal{E}_t^{fU}(\mathcal{L}_1) \subseteq \mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_2)$ si et seulement si

$\mathcal{E}_t^{fU}(\mathcal{L}_1) \subseteq \mathcal{L}_2$

et $\mathcal{E}_t^{U}(\mathcal{L}_1) \subseteq \mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_2)$ si et seulement si

$\mathcal{E}_t^{U}(\mathcal{L}_1) \subseteq \mathcal{L}_2$

2) Si $\mathcal{L}_1$ n'est pas fermé par t alors

$$\mathcal{E}_t^{fU}(\mathcal{L}_1) - \mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_1) \neq \emptyset$$

et

$$\mathcal{E}_t^{U}(\mathcal{L}_1) - \mathcal{E}_{\underline{\mathcal{L}}}^{U}(\mathcal{L}_1) \neq \emptyset$$

3) $\mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_1)$ est fermé par t si et seulement si

$\mathcal{L}_1 = \mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_1) = \mathcal{E}_t^{fU}(\mathcal{L}_1)$ et si $\mathcal{L}$ est un cône rationnel alors

$\mathcal{E}_{\underline{\mathcal{L}}}^{U}(\mathcal{L}_1)$ est fermé par t si $\mathcal{L}_1 = \mathcal{E}_{\underline{\mathcal{L}}}^{U}(\mathcal{L}_1) = \mathcal{E}_t^{U}(\mathcal{L}_1)$.

Démonstration

Pour tout $L \in \mathcal{E}_t^{fU}(\mathcal{L}_1)$ le langage $Ch(L) \in \mathcal{E}_t^{fU}(\mathcal{L}_1)$. Si donc $\mathcal{E}_t^{fU}(\mathcal{L}_1) \subseteq \mathcal{E}_{\underline{\mathcal{L}}}^{fU}(\mathcal{L}_2)$ par application répétée de la

proposition 6.5 il en résulte que $L \in \mathscr{L}_2$. Le point 2) résulte du point 1) avec $\mathscr{L}_2 = \mathscr{L}_1$. La preuve s'achève comme pour le théorème 6.1.

∎

Corollaire 6.5

Les opérations chevron, crochet et crochet ouvert sont indépendantes des opérations substitution dans un cône rationnel $\mathscr{L}$ fortement 1-itératif.

Dans le cas de cônes rationnels quelconques le corollaire 6.5 n'est pas nécessairement vérifié. En effet si l'on considère la substitution dans les langages linéaires, notée Lin, on a trivialement :

$C(\underline{Rat}) \subseteq \mathscr{E}^{U}_{\underline{Lin}}(Rat)$ mais $C(Rat) = Lin$ n'est pas inclus dans Rat.

proposition 6.5 il en résulte que $L \in \mathscr{L}_2$. Le point 2) résulte du point 1) avec $\mathscr{L}_2 = \mathscr{L}_1$. La preuve s'achève comme pour le théorème 6.1.

■

Corollaire 6.5

Les opérations chevron, crochet et crochet ouvert sont indépendantes des opérations substitution dans un cône rationnel $\mathscr{L}$ fortement 1-itératif.

Dans le cas de cônes rationnels quelconques le corollaire 6.5 n'est pas nécessairement vérifié. En effet si l'on considère la substitution dans les langages linéaires, notée Lin, on a trivialement :

$C(\underline{Rat}) \subseteq \mathscr{L}^{U}_{\underline{Lin}}(Rat)$ mais $C(Rat) = Lin$ n'est pas inclus dans Rat.

CHAPITRE 7

LANGAGES VÉRIFIABLES
ET
OPÉRATIONS SYNTAXIQUES CLASSIQUES

A) INTRODUCTION

B) LANGAGES VERIFIABLES, PROPRIETES ELEMENTAIRES

C) CARACTERISATION ET COMPARAISON DE DIVERSES HIERARCHIES CLASSIQUES

D) SUR DIVERSES AUTRES FAMILLES DE LANGAGES

A) INTRODUCTION

Les langages vérifiables ou "checking automata languages" ont été introduits par Ginsburg, Greibach et Harrison dans [GGH] comme cas particulier des "one-way stack automata". Intuitivement ces automates unilatères peuvent écrire un mot en mémoire puis en consulter le contenu mais sans jamais plus le modifier. Ces langages ont été plus précisément étudiés dans [Gr5,Rod1,Rod6,Si1,Si2,Klin] où diverses hiérarchies sont présentées. Cette classe de langages se révèle particulièrement intéressante par plusieurs aspects. D'une part elle forme un cône rationnel (fidèle) qui est une FAL (fidèle). Cette analogie avec les langages algébriques permet d'y introduire des hiérarchies analogues à celles considérées dans les algébriques (cf [Gr1] et [Rod6]). D'autre part nous constatons que cette famille est fermée par toutes les opérations classiques sur les langages considérées à ce jour : substitution, homomorphic replication, crochet, crochet ouvert, etc... Ceci lui confère donc une structure algébrique beaucoup plus puissante que celle des langages algébriques qui justifie totalement l'étude que l'on peut en faire. Nous pensons que la plupart des conjectures sur les langages algébriques seront résolues une fois que leur équivalent dans les langages vérifiables le sera. Si les langages manipulés sont souvent moins "simples", nous disposons d'outils, tels que les opérations sur les langages, plus puissants que ceux classiquement utilisés jusque là, qui devraient nous le permettre.

Dans ce chapitre nous rappelons tout d'abord la définition des langages vérifiables puis nous caractérisons et comparons diverses hiérarchies en utilisant l'opération crochet ouvert et ses généralisations.

Ceci nous permet, en utilisant les résultats des chapitres précédents, d'améliorer un certain nombre de résultats de [Lat4] et [Gr8].

Nous achevons cette étude par la comparaison des hiérarchies avec un certain nombre de langages vérifiables et non vérifiables définis dans [Lat4] et [Gr8].

B) LANGAGES VERIFIABLES, PROPRIETES ELEMENTAIRES

Définition 7.1

Un accepteur vérificateur (AV en abrégé) est un 7-uple

$$M = (Q,X,\Gamma,C,\delta,q_0,F) \text{ où}$$

- Q est un ensemble fini d'états, $q_0 \in Q$ est l'état initial, $F \subseteq Q$ est l'ensemble des états finals.
- X est l'alphabet fini d'entrée.
- Γ est l'alphabet fini de mémoire disjoint de $\mathbb{N}$.
- C est un compteur dont le contenu appartient à $\mathbb{N}$.
- δ est une application de $Q \times (X \cup \{\varepsilon\}) \times (\Gamma^* \times C)$ dans les parties finies de $Q \times (\Gamma^* \cup \{0,1,-1\})$ telle que pour tout $q \in Q$, $a \in X \cup \{\varepsilon\}$, $\gamma \in \Gamma^*$ et $i \leq \mathbb{N}$, $1 \leq i \leq |\gamma|$ on a :

$$\delta(q,a,(\gamma,0)) \cap Q \times \{0,-1\} = \emptyset$$

$$\delta(q,a,(\gamma,i)) \cap Q \times \Gamma^* = \emptyset.$$

Définition 7.2

Une configuration de M est un triplet $(q,w,(y,i))$ tel que $q \in Q$, $w \in X^*$, $y \in \Gamma^*$, $0 \leq i \leq |y|$. Nous noterons $\vdash$ la relation entre configurations définie par :

a) Si $(q,aw,(\gamma y,0))$ est une configuration avec $q \in Q$, $a \in X \cup \{\varepsilon\}$, $w \in X^*$, $\gamma \in \Gamma^*$, $y \in \Gamma \cup \{\varepsilon\}$, alors :

- Si $(q',A) \in \delta(q,a,y)$, $q' \in Q$, $A \in \Gamma^*$ alors

$$(q,aw,(\gamma y,0)) \vdash (q',w,(\gamma yA,0))$$

- Si $(q',1) \in \delta(q,a,y)$ alors

$$(q,aw,(\gamma y,0)) \vdash (q',w,(\gamma y,1))$$

b) Si $(q,aw,(y_n \ldots y_2 y_1,i))$ est une configuration avec $q \in Q$, $a \in X \cup \{\varepsilon\}$, $w \in X^*$, $y_j \in \Gamma$, $1 \leq j \leq n$ et $1 \leq i \leq n$, alors :

- Si $(q',j) \in \delta(q,a,y_i)$ avec $j \in \{0,1,-1\}$, $1 \leq i+j \leq n$ alors $(q,aw,(y_n \ldots y_2 y_1,i)) \vdash (q',w,(y_n \ldots y_2 y_1,i+j))$. $\overset{*}{\vdash}$ désignera la fermeture réflexive et transitive de $\vdash$.

Intuitivement un AV est un accepteur unilatère muni d'un ruban mémoire dont l'accès est déterminé par le compteur C. Tant que C = 0 l'automate peut écrire en mémoire exactement comme un automate à pile et l'AV lit alors le "sommet" de la mémoire. Sitôt que C ≠ 0 alors l'automate ne peut plus modifier le contenu de sa mémoire. Si $y_n y_{n-1} \ldots y_2 y_1$, $y_i \in \Gamma$, $1 \leq i \leq n$, est alors le mot mémoire et si C = j, $1 \leq j \leq n$ alors l'AV lit le symbole y_j en mémoire.

Un mot w sera accepté par M si, partant de la configuration initiale $(q_0,w,(\varepsilon,0))$ l'automate parvient à un état terminal après avoir lu le mot w.

Formellement on a :

Définition 7.3

Soit $M = (Q,X,\Gamma,C,\delta,q_0,F)$ un AV, le langage accepté par M est

l'ensemble

$$L(M) = \{w \in X^* \mid \exists q_F \in F,\ y \in \Gamma^*,\ 1 \leq i \leq |y| :$$

$$(q_0,w,(\varepsilon,0)) \overset{*}{\vdash} (q_F,\varepsilon,(y,i))\}.$$

Un langage L est vérifiable s'il existe un AV M tel que L = L(M). Vérif désignera la famille des langages vérifiables.

On a alors le théorème fondamental :

Théorème 7.1

Vérif est une FAL principale fermée par substitution, chevron, crochet, crochet ouvert et Shuf de degré i, i ≥ 1, chevron et crochet ouvert d'ordre n, n ≥ 1 et "homomorphic replication".

Démonstration

Greibach établit dans [Gr5] que Vérif est une FAL principale fermée par substitution et dans [Gr3] par "homomorphic replication". Montrons, par exemple, que Vérif est fermé par crochet ouvert d'ordre n, les autres cas se traitant de façon analogue. Soient donc $L_1,\ldots,L_n$ des langages vérifiables, X = {a,b} un alphabet, $c_1,c_2,\ldots,c_n$ des nouveaux symboles. On sait [Rod1] que le langage $L = \{wc_1wc_2\ldots wc_nw \mid w \in \{a,b\}^*\}$ est vérifiable. Soit alors σ la substitution définie par σ(a) = a, σ(b) = b et $\sigma(c_i) = L_i$, $1 \leq i \leq n$. Alors $COn(L_1,\ldots,L_n) = \sigma(L)$ et donc Vérif est fermé par COn puisque fermé par substitution.

■

Du chapitre 2 il résulte alors trivialement que si L est un générateur de la FAL Vérif on a également Vérif = $\mathcal{E}(L)$. Les générateurs de Vérif ont une structure relativement complexe. Citons celui donné dans [GG3] :

Définition 7.4

Soit $X = \{a,b\}$ et $\overline{X} = \{\overline{a},\overline{b}\}$. Soit $\equiv$ la relation sur $(X \cup \overline{X})^*$ définie par $\mu(a\overline{a}) \equiv \mu(b\overline{b}) \equiv \varepsilon$. Pour tout mot $x \in (X \cup \overline{X})^*$, $\mu(x)$ désignera le mot de longueur minimale tel que $x \equiv \mu(x)$.

Soit alors le langage

$$L = \{wcy \mid w \in X^*,\ y \in (X \cup \overline{X})^*,\ \mu(y) = \varepsilon,$$

$$\mu(\mathrm{Init}(y)) \subseteq \mathrm{Init}(\tilde{w})\} \cup \{\varepsilon\}.$$

L est un générateur de Vérif [GG3].

C) CARACTERISATION ET COMPARAISON DE DIVERSES HIERARCHIES CLASSIQUES

Plusieurs hiérarchies de langages vérifiables ont été considérées ces dernières années. Ces hiérarchies étaient soit définies à partir du fonctionnement des accepteurs vérificateurs [Rod1,Rod6,Si2] de manière analogue à [Gr1] pour les automates à pile, soit définies à partir de grammaires particulières [Si1].

Dans [Gr3] S. Greibach signale que certaines de ces familles de langages peuvent être caractérisées par les opérations syntactiques binaires "homomorphic replication" (langages reconnus par des accepteurs vérificateurs à mouvements bornés [Rod2], et par la "duplication" ("equal matrix languages" de [Si1])).

Dans [Klin] Klingenstein introduit une nouvelle classe de langages appelés "ρ" matrix languages" dont il donne une triple caractérisation : grammaticale ("ρ matrix grammars"), algébrique ("homomorphic replication") et en termes d'automates dans un cas particulier ("ρ checking automata"). Enfin dans [Gr8], considérant, sous un aspect complexité de calcul, la notion de machine à mouvements bornés notamment utilisée dans [BB,GS2,Gr1,Rod2,BNP], S. Greibach définit les "one-way preset Türing machines" qui lui permettent de retrouver des résultats des papiers cités ci-dessus et d'établir les liens existant entre plusieurs hiérarchies.

Dans ce paragraphe nous reprenons cette étude en utilisant les opérations chevron et crochet ouvert et leurs généralisations, ce qui nous permet d'améliorer des résultats de [Lat4] et [Gr8].

Définition 7.5

Soit $M=(Q,X,\Gamma,C,\delta,q_0,F)$ un accepteur vérificateur et $C_0,C_1,\ldots,C_p,\ldots,C_r$ une suite de configurations de M telles que $C_0 = (q_0,w,(\varepsilon,0)),C_{p-1} = (q_{p-1},w_{p-1},(y,0)),C_p = (q_p,w_p,(y',1))$ et $C_r = (q_F,\varepsilon,(y',i))$ avec $q_F \in F$, $y, y' \in \Gamma^*$, $1 \leq i \leq |y'|$. Pour tout j,

$1 \leq j \leq |y'|$ le nombre de consultation de la case mémoire j est le nombre de fois que le compteur de M prend la valeur j. L'analyse du mot w par M est à consultation k-bornée si le nombre de consultation de chaque case mémoire est au plus égal à k.

Un mouvement se produit en C_i si l'état du compteur dans la configuration C_i est différent de celui de C_{i+1}, s'il existe $j \leq i-1$ tel que le compteur de C_i est égal au compteur de C_u pour $j+1 \leq u \leq i$ et le compteur de C_{i+1} est égal à celui de C_j. Le nombre de mouvements d'une analyse est $1+\text{card}\{i \mid \text{un mouvement se produit en } C_i\}$. Une analyse est à mouvement k-borné si le nombre de mouvements est inférieur ou égal à k.

Intuitivement M fait un mouvement si son compteur passe de l'incrémentation à la décrémentation ou inversement.

Définition 7.6

Un accepteur vérificateur M est à consultation k-bornée (resp. à mouvement k-borné) si pour tout $w \in L(M)$ il existe une analyse de w par M qui est à consultation (resp. à mouvement) k-borné.

Cons_k (resp. Mouv_k) désignera la famille des langages reconnus par un accepteur vérificateur à consultation (resp. mouvement) k-borné. Nous poserons enfin $\text{Cons} = \bigcup_{k \geq 1} \text{Cons}_k$ et $\text{Mouv} = \bigcup_{k \geq 1} \text{Mouv}_k$.

Afin de comparer les diverses hiérarchies classiques de langages vérifiables, rappelons la définition de l'opération "duplication" introduite dans [Gr3] comme cas particulier de "l'homomorphic replication".

Définition 7.7

L'opération "homomorphic replication" d'ordre n ρ est une duplication d'ordre n si $\rho(i) = 1$, $1 \leq i \leq n$.

On a alors

Proposition 7.1 [Klin,Gr3]

Pour tout $n \geq 1$, $\mathfrak{D}_n$, la famille des "equal matrix languages" d'ordre n ou "n-turn checking automata languages" [Si1,Si2] est égal au cône rationnel engendré par la duplication d'ordre n des langages rationnels.

La famille $\mathfrak{D} = \bigcup_{n \geq 1} \mathfrak{D}_n$ des "equal matrix languages" est la fermeture par duplication des langages rationnels.

Proposition 7.2 [Gr3,Gr8]

Pour tout $k \geq 1$, la famille Mouv_k est le cône rationnel engendré par "l'homomorphic replication" d'ordre k des langages rationnels.

$\text{Mouv} = \bigcup_{k \geq 1} \text{Mouv}_k$ est la fermeture par "homomorphic replication" des langages rationnels.

Nous pouvons donner une autre caractérisation de ces familles respectivement par le crochet ouvert d'ordre n et une opération syntactique unaire introduite par Latteux dans [Lat4] :

Définition 7.8

Pour tout langage L et tout $k \geq 1$, soit {a,b} un alphabet disjoint de X_L. Nous poserons

$$(L)_{2k} = \{wx_1\tilde{w}x_2w\ldots wx_k\tilde{w} \mid w \in \{a,b\}^*,\ x_i \in L,\ 1 \leq i \leq k\}$$

et

$$(L)_{2k-1} = \{wx_1\tilde{w}\ldots wx_k \text{ et } x_1w\ldots\tilde{w}x_kw \mid w \in \{a,b\}^*,\ x_i \in L,\ 1 \leq i \leq k\}$$

$(\ )_k$ désignera également l'opérateur associé à $(\ )_k$. On a alors les deux

propositions

Proposition 7.3

Pour tout $n \geq 1$, COn(Rat) est égal à $\mathcal{D}_{n+1}$, la famille des "equal matrix languages" d'ordre n+1, et donc $\mathcal{D} = \bigcup_{n \geq 1} CO_n(Rat)$.

Proposition 7.4

Pour tout $k \geq 1$, la famille $Mouv_k$ est le cône rationnel $(Rat)_{2k}$. Mouv est donc égal à $\bigcup_{k \geq 1} (Rat)_{2k}$.

Nous allons maintenant comparer ces diverses familles. On a tout d'abord :

Lemme 7.1

Pour tout $n \geq 1$, la famille $\mathcal{D}_3$ n'est pas incluse dans $CO^n(Rat)$.

Démonstration

Rat est une famille de langages fortement 1-itérative. D'après la proposition 6.3 il en résulte que pour tout $n \geq 1$, $CO^n(Rat)$ est fortement 2-itératif. Or $\mathcal{D}_3$ est le cône rationnel engendré par $K_2 = \{(wc)^2 w | w \in \{a,b\}^*\}$ et ce langage n'est pas fortement 2-itératif.

■

D'où :

Proposition 7.5

1) $CO(Rat) = \mathcal{D}_2$.

2) Pour tout $n \geq 2$, $CO^n(Rat) \neq \mathcal{D}_{2n}$.

Démonstration

On a clairement $CO(Rat) = \mathcal{D}_2 = \mathcal{E}(\{wcw|w \in \{a,b\}^*\})$. Pour montrer que $CO^n(Rat) \subseteq \mathcal{D}_{2n}$, il suffit de vérifier que pour tout $n \geq 2$, le langage $L_n = \{w_1cw_2c\ldots w_ndw_n\ldots cw_2cw_1|w_i \in \{a,b\}^*, 1 \leq i \leq n\}$ appartient à $\mathcal{D}_{2n}$. Soit donc $\{a_1,\bar{a}_1,\ldots,a_n,\bar{a}_n\}$ un nouvel alphabet et $A_n = \{x_1x_2\ldots x_nd|x_i \in \{a_i,\bar{a}_i\}^*, 1 \leq i \leq n\}$. A_n est un langage rationnel. Soit alors ρ la duplication d'ordre 2n définie par $\varphi_i(x) = x$ si $x \in \{a_i,\bar{a}_i\}$, $\varphi_i(x) = \varepsilon$ sinon, $1 \leq i \leq n$, $\varphi_n(d) = d$ et $\varphi_{n+i} = \varphi_{n-i+1}$ pour $1 \leq i \leq n$. Alors, par définition

$$M_n = \{\varphi_1(w)\ldots\ \varphi_n(w)\ \varphi_{n+1}(w)\ldots\ \varphi_{2n}(w)|w \in A_n\}$$

est égal à

$$\{x_1x_2\ldots x_ndx_n\ldots x_2x_1|x_i \in \{a_i,\bar{a}_i\}^*, 1 \leq i \leq n\}$$

et appartient à $\mathcal{D}_n$. Or L_n est rationnellement équivalent à M_n, d'où le résultat.

∎

Remarque

On peut démontrer, en utilisant notamment des résultats de [Lat4], que $CO^n(Rat)$ est incomparable à $\mathcal{D}_n$. De la proposition précédente, il résulte donc

Corollaire 7.1

$\mathcal{E}^U_{CO}(Rat)$ est strictement inclus dans la famille des "equal matrix languages".

Proposition 7.6 [Klin]

Pour tout $k \geq 1$, la famille $\mathcal{D}_k$ est strictement incluse dans la famille Mouv_k. Il en résulte que $\mathcal{D} \subsetneq \text{Mouv}$.

Démonstration

L'inclusion $\mathcal{D}_k \subseteq \text{Mouv}_k$ est évidente pour tout k. L'inclusion stricte se démontre sans peine à l'aide d'un lemme analogue au lemme 6.2 pour l'opération crochet ouvert d'ordre n.

■

Nous achèverons cette étude par la comparaison des familles Mouv_k et Cons_k, $k \geq 1$. Rappelons tout d'abord les résultats suivants de [Rod6,Gr8] :

Définition 7.9

Soit $X = \{a,b\}$ et $\overline{X} = \{\overline{a},\overline{b}\}$. Pour tout $w \in (X \cup \overline{X})^*$ soit $\mu(w)$ l'élément de longueur minimale de la classe de w modulo la congruence $a\overline{a} \equiv b\overline{b} \equiv \varepsilon$. Pour tout $L \in X^*$, soit $c(L) = \{wcy \mid w \in L,\ y \in (X \cup \overline{X})^*,\ \mu(\text{Init}(y)) \subseteq \text{Init}(w)\}$ et pour tout $k \geq 1$

$$c_k(L) = \{wcy \in c(L) \mid \text{card}\{y' \in \text{Init}(y) \mid |\mu(y')| = i\} \leq k,\ 1 \leq i \leq |w|\}$$

On a alors

Théorème 7.2 [Rod6]

Pour tout $k \geq 1$, la famille Cons_k est égale au cône rationnel engendré par $c_k(\{a,b\}^*)$. De plus on a $\text{Cons}_k \subsetneq \text{Cons}_{k+1}$.

Nous retrouvons alors sans peine le

Théorème 7.3 [Gr8]

Pour tout $k \geq 3$:

1) $Mouv_k \not\subseteq Cons_k$.

2) $Cons_3$ et $Mouv_{k+1}$ sont incomparables ainsi que $Cons_3$ et Mouv.

Démonstration

1) L'inclusion $Mouv_k \subseteq Cons_k$ est évidente. D'autre part $Cons_k$, $k \geq 3$, contient S_2^* mais n'appartient pas à $Mouv_k = CO_k(Rat)$ d'après la proposition 5.2 et le corollaire 6.3 ce qui prouve le point 1).

2) D'autre part le langage $K_2 = \{(wc)^2 w | w \in \{a,b\}^*\}$ appartient à $Mouv_{k+1}$, $k \geq 3$ mais n'appartient pas à $Cons_3$ comme on peut le montrer sans difficulté par des moyens analogues à ceux de [Gr1,Rod6 ou Gr8].

■

D) SUR DIVERSES AUTRES FAMILLES DE LANGAGES

Dans [Lat4] et [BGW] les auteurs étudient un certain nombre de cônes rationnels (fidèles) et FAL (fidèles) qui sont définis soit à partir de machines particulières [BGW] soit à partir d'ETOL-Systèmes (cf [Roz]) dans [Lat4]. Nous reprenons et complétons ici ces études d'un point de vue algébrique en utilisant les opérations définies dans les chapitres précédents.

Nous utiliserons sans les rappeler formellement les définitions de [BGW] et de [Lat4]. Nous renvoyons le lecteur à ces articles pour plus de précision.

Intuitivement une "single reset machine" est un accepteur vérificateur qui fait au plus un mouvement et qui fonctionne en temps "presque réel", i.e. il existe un entier k tel que l'AV effectue au plus k transitions consécutives en lisant le mot vide.

On a alors :

Théorème 7.4 [BGW]

La famille des langages acceptés par les "single reset machines" est exactement le cône rationnel fidèle engendré par $L_1 = \{wcw | w \in \{a,b\}^*\}$.

De l'étude du crochet ouvert faite au chapitre 3 il en résulte immédiatement :

Corollaire 7.2 [BGW]

Un langage L appartient à la famille des langages acceptés par les "single reset machines" si et seulement si L appartient à $\mathcal{E}^f(L_1) = \mathcal{E}(L_1)$ ou si et seulement si il existe un langage rationnel R et deux homomorphismes φ_1 et φ_2 tels que $L = \{\varphi_1(w)\ \varphi_2(w) | w \in R\}$.

Contrairement à [BGW] ce corollaire est établi d'un point de vue purement algébrique, sans considération de machines. De la même façon dans [BGW] la famille des langages acceptés par des "reusable reset tape machines" fonctionnant en temps presque réels est identifiée au cône rationnel fidèle engendré par $(L_1c)^*$. Il résulte alors directement de notre étude que cette famille est aussi égale au cône rationnel et à la FAL (fidèle) engendrée par $(L_1c)^*$.

Plus intéressant est le cas du langage $*COPY = \{(wc)^n | n \geq 0, w \in \{a,b\}^*\}$ encore appelé R_∞ dans [Lat4]. Dans [BGW] les auteurs démontrent que ce langage est un générateur fidèle du cône rationnel fidèle des langages reconnus par les "one non writing circular tape machines" fonctionnant en temps presque réel. Latteux établit qu'il s'agit de la famille NLID(Rat) obtenue en remplaçant l'axiome ω d'un ETOL-Système non déterministe linéaire à droite par un langage rationnel (cf [Lat4]).

Nous allons voir que cette famille peut être caractérisée à l'aide d'une opération unaire, notée $CO\infty$, qui peut être considérée comme la limite des opérations crochet ouvert d'ordre n sous forme unaire :

Définition 7.10

Soit $CO\infty$ l'opération unaire sur les langages, définie pour tout langage $L \subseteq X^*$ si $\{a,b\}$ est un alphabet disjoint de X par :

$$CO\infty(L) = \{w\ell_1w\ell_2\ldots w\ell_n | w \in \{a,b\}^*, \ell_i \in L, 1 \leq i \leq n\}.$$

Comme l'opération crochet est indépendante de Rat on a

Proposition 7.7

$CO\infty$ est une opération indépendante de Rat.

Définition 7.11

Pour tout cône rationnel $\mathcal{L}$, on désignera par $CO\infty(\mathcal{L})$ l'ensemble

$$CO\infty(\mathcal{L}) = \mathcal{E}^U(\{CO\infty(L) | L \in \mathcal{L}\})$$

et par

$$CO\infty^f(\mathcal{L}) = \mathcal{E}^{fU}(\{CO\infty(L) | L \in \mathcal{L}\}).$$

D'après [BGW] on a alors

Proposition 7.8 [BGW] [Lat4]

$$CO\infty^f(Rat) = \mathcal{E}^f(*COPY) \subsetneq CO\infty(Rat) = \mathcal{E}(*COPY)$$

$$\subsetneq \underline{Rat}_0\ CO\infty^f(Rat) = \mathcal{P}^f(*COPY) \subsetneq \underline{Rat}\ CO\infty(Rat) = \mathcal{P}(*COPY)$$

$\subsetneq$ RE la famille des langages récursivement énumérables.

Ceci montre donc que si $\{t_i\}$ est une suite d'opérations indépendante de Rat alors sa limite est indépendante de Rat. Par contre si les t_i sont de degré p alors la limite ne l'est pas nécessairement (cf le crochet ouvert de degré infini). De même si les t_i préserve l'effacement, il n'en est pas nécessairement de même pour la limite.

Rappelons la proposition de [Lat4].

Proposition 7.9 [Lat4]

Les langages $S_2 = \{wc\tilde{w} | w \in \{a,b\}^*\}$ et

$$L_\infty = \{w_n c w_{n-1} c \ldots w_1 d w_1 \ldots c w_{n-1} c w_n | n \in \mathbb{N}, w_i \in \{a,b\}^*,\ 1 \leq i \leq n\}$$

n'appartient pas à $CO\infty(Rat) = \mathcal{E}(*COPY)$.

Comme Ch est indépendant de Rat et que S_2 et L_∞ sont respectivement rationnellement équivalents à $Ch(S_2)$ et $Ch(L_\infty)$, il en résulte :

Corollaire 7.3

Les langages S_2 et L_∞ n'appartiennent pas à la FAL engendrée par *COPY.

Réciproquement le langage $\mathcal{E}(L_\infty)$ est fortement 2-itératif tandis que *COPY ne l'est pas, on a donc

Théorème 7.5

Les cônes rationnels $\mathcal{E}(L_\infty)$ et $\mathcal{E}(*COPY)$ sont incomparables. Les FAL $\mathcal{F}(L_\infty)$ et $\mathcal{F}(*COPY)$ sont incomparables.

Démonstration

Le premier point résulte directement des considérations précédentes. $L_\infty \notin \mathcal{F}(*COPY)$ d'après le corollaire 7.2 et *COPY n'appartient pas à $\mathcal{F}(L_\infty) = \underline{Rat}\ \mathcal{E}(L_\infty)$ pour une raison semblable puisque *COPY est équivalent à Ch(*COPY) d'après [Lat4].

■

Nous terminerons cette étude en considérant le langage "Shift" de [BGW] :

Définition 7.12

Soient x et y deux mots de X^* de même longueur ; si $x = x_1x_2\ldots x_n$ et $y = y_1y_2\ldots y_n$, x_i et $y_i \in X$, $1 \leq i \leq n$ nous désignerons par $\langle x,y \rangle$ le mot de $(X \times X)^*$ défini par $\langle x,y \rangle = (x_1,y_1)(x_2,y_2)\ldots(x_n,y_n)$.

Si c est un nouveau symbole n'appartenant pas à X = {a,b}, nous appellerons Shift le langage

$$\text{Shift} = \{ \langle w_1,w_2\rangle \; c \; \langle w_2,w_3\rangle \; c \ldots \langle w_{m-1},w_m\rangle \; c \mid m \leq 2, \; w_i \in X^*, |w_1| = \ldots = |w_n| \} \cup \{\varepsilon\}.$$

On a

Théorème 7.6 [BGW]

$\mathcal{E}^f(\text{Shift}) \subsetneq \mathcal{F}^f(\text{Shift}) \subsetneq \mathcal{E}(\text{Shift}) = \mathcal{F}(\text{Shift}) = \text{RE}$ la famille des langages récursivement énumérables.

$\mathcal{E}^f(\text{Shift})$ est en outre la famille des langages acceptés par les "one-writing circular tape machines" fonctionnant en temps presque réel [BGW].

Comme $\mathcal{E}(\text{Shift})$ est égal à la famille des langages récursivement énumérables il en résulte trivialement que Shift n'est pas un langage vérifiable et n'appartient donc à aucune des familles $\mathcal{E}^U_{CO}(\text{Rat})$, $\mathcal{E}^U_{CO_n}(\text{Rat})$, ni $\mathcal{E}^U_{CO_n}(\text{Rat})$, $n \geq 1$. Réciproquement ces familles étant récursivement énumérables appartiennent à $\mathcal{E}(\text{Shift})$. Par contre d'après [BGW, théorème 4.1] on a :

Théorème 7.7 [BGW]

1) $\mathcal{E}^f(\text{Shift})$ est incomparable à $\mathcal{E}^f(L_1)$ et $\mathcal{F}^f(L_1)$.

2) $\mathcal{E}^f((L_1c)^*) \subsetneq \mathcal{E}^f(\text{Shift})$.

3) $\mathcal{E}^f(\text{Shift})$ est incomparable à $\mathcal{F}^f(*\text{COPY})$, $\mathcal{E}^f(*\text{COPY})$ et $\mathcal{E}(*\text{COPY})$.

4) $\mathcal{F}^f(*\text{COPY}) \subsetneq \mathcal{F}^f(\text{Shift})$.

5) $\mathcal{F}^f(\text{SHIFT})$ est incomparable à $\mathcal{E}(*\text{COPY})$ et $\mathcal{F}(*\text{COPY})$.

BIBLIOGRAPHIE

[Aut] AUTEBERT J.M.
Non principalité du cylindre des langages à compteur
Mathem. Systems Theory 11 (1977), 157-167.

[Aut2] AUTEBERT J.M.
Cylindres de langages simples et pseudo-simples
3rd GI Conf. on Theoretical Computer Science, Lecture Notes in Comp. Science, 48 (1977), 149-153.

[Aut3] AUTEBERT J.M.
Cylindres de langages algébriques
Thèse de Doctorat d'Etat, Université de Paris 6, Octobre 1978.

[BB] BAKER B., BOOK R.V.
Reversal-bounded multi-pushdown machines
Journ. of Comp. and System Science 8 (1974), 315-332.

[Beau1] BEAUQUIER G.
A Remark about the syntactic lemma
A paraître.

[Beau2] BEAUQUIER G.
Contribution à l'étude de la complexité structurelle des langages algébriques
Thèse de doctorat d'Etat, Université de Paris 6, Mars 1977.

[Berst] BERSTEL J.
Une hiérarchie des parties rationnelles de $\mathbb{N}^2$
Math. Systems Theory 7 (1973), 114-137.

[Bo1] BOASSON L.
Paires itérantes et langages algébriques
Thèse de doctorat d'état, Université de Paris 7, 1974.

[Bo2] BOASSON L.
Cônes rationnels et familles agréables de langages
Application aux langages à compteur
Thèse 3ème cycle, Université de Paris 7, 1971.

[BCoN] BOASSON L., COURCELLE B., NIVAT M.
A new complexity measure for languages
in Proceedings of a conference on theoretical computer Science Waterloo (1977), 130-138.

[BCN] BOASSON L., CRESTIN J.P., NIVAT M.
Familles de langages translatables et fermées par crochet
Acta Informatica 2 (1973), 383-393.

[BoNi1] BOASSON L., NIVAT M.
Sur diverses familles de langages fermées par transduction rationnelle
Acta Informatica 2 (1973), 180-188.

[BG] BOOK R.V., GRAIBACH S.
The independence of certain operations on semi AFLs
RAIRO Inform. Théorique, vol. 12, n°4 (1978), 369-385.

[BGW] BOOK R.V., GREIBACH S., WRATHALL C.
Comparisons and reset machines
Automata, languages and Programming 5th. Colloquium Udine (1978) Lecture Notes in Computer Science 62 (1978), 113-124.

[BN] BOOK R.V., NIVAT M.
Linear languages and the intersection closures of classes of languages
SIAM Journ. Comput. 7 (1978), 167-177.

[BNP] BOOK R.V., NIVAT M., PATERSON M.
Reversal-bounded acceptors and intersections of linear languages
SIAM Journ. Comput. 3 (1974), 283-295.

[Chof] CHOFFRUT C.
Contribution à l'étude de quelques familles remarquables de fonctions rationnelles
Thèse de doctorat d'état, Université Paris 7, Novembre 1978.

[Chom1] CHOMSKY N.
Three models for the description of languages
IRE Trans. Information Theory 112 (1956), 113-124.

[Chom2] CHOMSKY N.
On certain formal properties of grammars
Info. and Control 2 (1959), 137-167.

[CS] CHOMSKY N., SCHUTZENBERGER M.P.
The algebraic theory of context-free languages
Computer Programming and Formal Systems, North Holland Pub. (1963).

[Cour] COURCELLE B.
Recursive Shemes, algebraic trees and deterministic languages
Proc. of 15 th Annual Symposium on Switching and Automata Theory (1974), 52-62.

[CR] COUSINEAU F.G., RIFFLET J.M.
Langages d'interprétation des schémas récursifs
RAIRO R1 (1975), 21-42.

[Crest] CRESTIN J.P.
Sur un langage quasi-rationnel d'ambiguité non bornée
Thèse 3ème cycle, Fac. Sciences de Toulouse (1967).

[Crest2] CRESTIN J.P.
Langages quasirationnels, ultralinéaires et bornés
Rapport de recherche 052, ENSTA (1975).

[Dur] DURIEUX J.L.
Sur l'image par une transduction rationnelle des mots sur une lettre
RAIRO R-2 (1975), 25-37.

[Eil] EILENBERG S.
Communication au congrès international des mathématiciens, Nice (1970).

[Ei2] EILENBERG S.
Automata, languages and machines, vol. A, Acad. Press (1974).

[EM] ELGOT C.C., MEZEI J.F.
On relations defined by generalized finite automata
IBM J. of Res. and Dev. 9 (1962), 47-68.

[Flie] FLIESS M.
Sur certaines familles de séries formelles
Thèse de doctorat d'état, Paris 7 (1972).

[Gi1] GINSBURG S.
The mathematical theory of context-free languages
Mac Graw Hill (1966).

[Gi2] GINSBURG S.
Algebraic and automata-theoretic properties of formal languages
North Holland, Pub. Comp. (1975).

[GGG] GINSBURG S., GOLDSTINE J., GREIBACH S.
Uniformly erasable AFL
Journ. of Comput. and Syst. Sciences 10 (1975), 165-182.

[GG1] GINSBURG S., GREIBACH S.
Abstract families of languages
Mem. of Amer. Math. Soc. 87 (1969), 1-32.

[GG2] GINSBURG S., GREIBACH S.
Principal AFL
Journ. of Comput. and Syst. Sciences 4 (1970), 308-338.

[GG3] GINSBURG S., GREIBACH S.
On AFL generators for finitely encoded AFA
Journ. of Comput. and Syst. Sciences 7 (1973), 1-27.

[GG4] GINSBURG S., GREIBACH S.
Deterministic context-free languages
Inform. and Control 9 (1966), 620-648.

[GGH1] GINSBURG S., GREIBACH S., HARRISON M.
One-way stack automata
Journ. Ass. Comp. Mach. 14 (1967), 389-418.

[GGH2] GINSBURG S., GREIBACH S., HARRISON M.
Stack automata and compiling
Journ. Ass. Comp. Mach. 14 (1967), 172-201.

[GHa] GINSBURG S., HARRISON M.
On the closure of AFL under reversal
Info. and Control 17 (1970), 395-409.

[GS1] GINSBURG S., SPANIER E.H.
Substitutions in families of languages
Inform. Sciences 2 (1970), 83-110.

[GS2] GINSBURG S., SPANIER E.H.
Finite-turn pushdown automata
SIAM Journ. Control (1966), 429-453.

[GS3] GINSBURG S., SPANIER E.H.
AFL with the semilinear property
Journ of Compu. and Syst. Sciences 5 (1971), 365-396.

[Gold1] GOLDSTINE J.
Abstract families of languages generated by bounded languages
SDC Doc. TM-738/059/00, Santa Monica (1970).

[Gold2] GOLDSTINE J.
Substitution and bounded languages
Journ. of Comput. and Syst. Sciences 6 (1972), 9-29.

[Gr1] GREIBACH S.
An infinite hierarchy of context-free languages
Journ. Ass. Comp. Mach. 16 (1969), 91-106.

[Gr2] GREIBACH S.
Chains of full AFLs
Math. Syst. Theory 4 (1970), 231-242.

[Gr3] GREIBACH S.
Syntactic operators on full semi AFLs
Journ. of Comput. and Syst. Sciences 6 (1972), 30-76.

[Gr4] GREIBACH S.
The unsolvability of the recognition of linear context-free languages
Journ. Ass. Comp. Mach. 13 (1966), 582-587.

[Gr5] GREIBACH S.
Checking automata and one-way stack languages
Jour. of Comput. and Syst. Sciences 3 (1969), 196-217.

[Gr6] GREIBACH S.
Independence of substitution and homomorphic replication
RAIRO Inform. Theorique, vol. 12, n°3 (1978), 213-234.

[Gr7] GREIBACH S.
Control sets on context-free grammar forms
Journ. of Comput. and Syst. Sciences 15 (1977), 35-98.

[Gr8] GREIBACH S.
One-way finite visit automata
Theo. Comp. Science 6 (1978), 175-222.

[Gr9] GREIBACH S.
Remarks on the complexity of non deterministic counter languages
Theo. Comp. Science 1 (1976), 269-288.

[Gr10] GREIBACH S.
A new normal form theorem for context-free languages
Journ. Ass. Comp. Mach. 16 (1969), 91-106.

[GH] GREIBACH S., HOPCROFT J.E.
Independence of AFL operations, in Studies in Abstract Families of Languages
Memoirs of the Amer. Math. Society 87 (1969), 33-40.

[Greu] GREUSSAY P.
Communication personnelle.

[HH] HARTMANIS J., HOPCROFT J.E.
What makes some language theory problems undecidable
Journ. of Comput. and Syst. Sciences 4 (1970), 368-376.

[Ito] ITOGA
Comparing language operations
Math. Syst. Theory 10 (1977), 305-321.

[Jac1] JACOB G.
Sur un théorème de Shamir
Info. and Control 27 (1975), 218-261.

[Jac2] JACOB G.
Un théorème de factorisation des produits d'endomorphismes de K^N
Journal of Algebra (1977).

[Klin] KLINGENSTEIN K.
ρ-matrix languages,
Theoretical Comp. Scien. - A paraître.

[Knu] KNUTH D.E.
One the translation of languages from left to right
Info. and Control 8 (1965), 607-639.

[Lat1] LATTEUX M.
Langages commutatifs
Thèse de doctorat d'état, Université de Lille (1978).

[Lat2] LATTEUX M.
Une note sur le lemme syntaxique de Greibach
Publi 106, Labo. Calcul de Lille (1978).

[Lat3] LATTEUX M.
Cônes rationnels commutativement clos
RAIRO Inform. Théorique 11 (1977), 29-51.

[Lat4] LATTEUX M.
EDTOL Systèmes ultralinéaires et opérateurs associés
Publi 100, Labo Calcul de Lille (1977).

[Mi] MINSKY M.L.
Recursive Unsolvability of Post's problem of "tag" and other topics in the theory of Türing machines
Annals of Mathem. 74 (1961), 437-455.

[Ni1] NIVAT M.
Transduction des langages de Chomsky
Ann. Inst. Fourieur (Grenoble) 18 (1968), 339-445.

[Ni2] NIVAT M.
Transduction des langages de Chomsky
Thèse de doctorat d'état, Université Paris (1967).

[Ogd] OGDEN W.F.
A helpful result in proving hinerent ambiguity
Math. Syst. Theory 2 (1968), 191-194.

[Rod1] RODRIGUEZ F.
Cônes d'accepteurs - Application à l'étude d'une hiérarchie infinie de cônes rationnels de langages d'accapteurs vérificateurs
Thèse de 3ème cycle, Université de Toulouse (1973).

[Rod2] RODRIGUEZ F.
Note sur la non principalité de certaines FAL
RAIRO Infor. Theorique 12 (1978), 181-184.

[Rod3] RODRIGUEZ F.
Cônes rationnels commutativement clos, d'après M. LATTEUX
Séminaire d'Informatique Théorique de l'Institut de Programmation Paris in "Actes du Séminaire d'Informatique Théorique Années 1975-1976 et 1976-1977, 90-105

[Rod4] RODRIGUEZ F.
Familles de langages fermées par crochet et crochet ouvert
3rd GI Conference on Theoretical Computer Science, Darmstadt 28-30 Mars 1977, Lecture Notes in Computer Science 48 (1977), 154-168.

[Rod5] RODRIGUEZ F.
Familles de langages fermées par crochet ouvert
A paraître dans Theorical Computer Science (1979).

[Rod6] RODRIGUEZ F.
Une double hiérarchie infinie de langages vérifiables
RAIRO Info. Theorique R1 (1975), 5-20.

[Rog] ROGERS H.
Theory of recursive functions and effective computability
Mac Graw-Hill (1967).

[Roz] ROZENBERG G.
Extension of tabled OL-Systems and languages
Intern. Jour. Comp. Inf. Sci. 2 (1973), 311-336.

[Schütz1] SCHUTZENBERGER M.P.
Certain elementary families of automata
Symposium on Math. Theory of Automata, Polytechnic Institue of Brookley (1962), 139-153.

[Schütz2] SCHUTZENBERGER M.P.
On context-free languages and pushdown automata
Inform. and Control 6 (1963), 246-264.

[Si1] SIROMONEY R.
On equal matrix languages
Inform. and control 14 (1969), 135-161.

[Si2] SIROMONEY R.
Finite-turn checking automata
Journ. of Comput. and Syst. Sciences 5 (1971), 549-559.

[VdW] VAN DER WALT A.
Locally linear families of languages
Inform. and Control 32 (1976), 27-32.